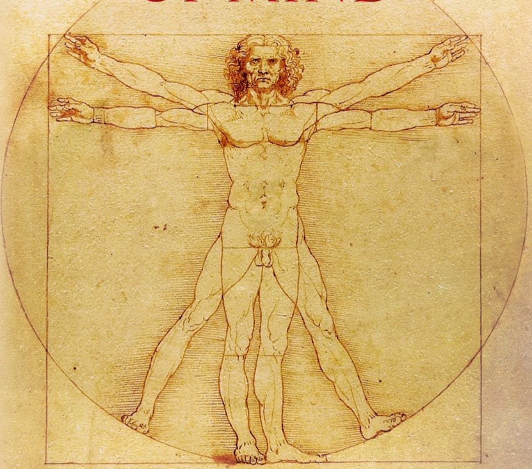
OXFORD

WILLIAM JAWORSKI

STRUCTURE AND THE METAPHYSICS OF MIND



How Hylomorphism Solves the Mind-Body Problem

Structure and the Metaphysics of Mind

Structure and the Metaphysics of Mind

How Hylomorphism Solves the Mind-Body Problem

William Jaworski

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2016

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2015955587

ISBN 978–0–19–874956–1

Printed in Great Britain by
Clays Ltd, St Ives plc

For my children

Contents

Acknowledgments

I'm grateful to numerous people for their help and feedback on material from the manuscript. They include Ken Aizawa, Ned Block, Godehard Brüntrup, Dave Chalmers, Joe Corabi, Bryan Frances, Carl Gillett, John Heil, Anna Marmodoro, Brian McLaughlin, Howard Robinson, William Simpson, Joe Vukov, Dean Zimmerman, two anonymous referees for OUP, and participants at the Fordham-Rutgers Metaphysics of Mind Conference, the Fordham Mind and Metaphysics Workshops, the Templeton Foundation's Analytic Theology Summer School held in July 2012 through the Hochschule für Philosophie München, and the Powers, Perception, and Agency Summer School and Conference held in September 2015 at the British School in Rome under the auspices of Corpus Christi College, Oxford, and the Power Structuralism in Ancient Ontologies project. I am grateful to Cornell University Press and Peter van Inwagen for their permission to reproduce sections of van Inwagen's *Material Beings*, and to Peter Momtchiloff, Eleanor Collins, and the staff at Oxford University Press for their unflagging support.

Some chapters include by permission versions of material that has appeared elsewhere. With kind permission from Wiley-Blackwell, Chapters 6, 10, 11, 12, 13, and 14 include versions of material that appeared in *Philosophy of Mind: A Comprehensive Introduction*, copyright © 2011 by William Jaworski. Chapters 1–4 and 6 include versions of material that appeared in 'Hylomorphism and the Metaphysics of Structure,' *Res Philosophica* 91(2014): 179–202, copyright © 2014 by William Jaworski. With kind permission of Taylor and Francis Group, LLC, a division of Informa plc (conveyed through Copyright Clearance Center, Inc.), Chapters 1, 6, 8, 10, and 13 include versions of material that appeared in 'Powers, Structures, and Minds,' from *Powers and Capacities in Philosophy: The New Aristotelianism*, edited by Ruth Groff and John Greco (New York: Routledge, 2013), 145–71, copyright © 2013 by William Jaworski. With kind permission from Springer Science+Business Media, Chapter 10 includes a version of material that appeared in 'The Logic of How-Questions,' *Synthese* 166 (2009): 133–55, copyright © 2009 by William Jaworski.

List of Figures

Introduction

> I regard the mind-body problem as wide open and extremely confusing.
>
> Saul Kripke

> [I]t is characteristic of our time, even while rejecting dualism, to take its questions to be fundamental starting points.
>
> Martha Nussbaum

> [T]he "solution" of the problem of mind-body is to be found in a revision of the preliminary assumptions . . . which generate the problem.
>
> John Dewey

> [W]ithin the framework of Aristotelian metaphysics and psychology there can be no mind-body problem.
>
> Mortimer Adler

> [H]ylomorphism is on the rise in contemporary metaphysics.
>
> Michael C. Rea[1]

The foregoing quotes suggest a research project, one that looks to solve mind-body problems by appeal to hylomorphism. Hylomorphism claims that structure (or organization, form, arrangement, order, or configuration) is a basic ontological and explanatory principle. Some individuals, paradigmatically living things, consist of materials that are structured or organized in various ways. You and I are not mere quantities of physical materials; we are quantities of physical materials with a certain organization or structure. That structure is responsible for us being and persisting as humans, and it is responsible for us having the particular developmental, metabolic, reproductive, perceptive, and cognitive capacities we have.

Hylomorphists take mind-body problems to be symptomatic of a worldview that rejects structure. Hylomorphic structure carves out distinctive individuals from the otherwise undifferentiated sea of matter and energy that is or will be described by our best physics, and it confers on those individuals distinctive powers. If hylomorphic structure exists, then the physical universe is punctuated with pockets of organized change and stability—composite physical objects (paradigmatically living things)

[1] Kripke 1972: 155; Nussbaum 1984: 207; Dewey 1958: 263; Adler 1993: 218; Rea 2011: 342.

whose structures confer on them powers that distinguish what they can do from what unstructured materials can do. If those powers include the abilities to think, feel, perceive, and act, then hylomorphic structure provides a way of locating mind in a physical world.

If structure is uncontroversially part of the physical world, and mental phenomena are species of structural phenomena, then they must be uncontroversially part of the physical world as well. A worldview that rejects hylomorphic structure, by contrast, is a worldview that lacks a basic principle which distinguishes the parts of the physical universe that can think, feel, and perceive from those that can't, and without a basic principle that carves out zones with distinctive powers, the existence of those powers in the natural world can start to look inexplicable and mysterious. If there is nothing built into the basic fabric of the universe that explains why Zone A has powers that Zone B lacks—if nothing explains why you, say, have the power to think, feel, and perceive, while the materials surrounding you do not, then the options for understanding the existence of those powers in the natural world become constrained: either those powers must be identified with the powers of physical materials taken by themselves or in combination (as panpsychists and many physicalists claim), or their existence must be taken as an inexplicable matter of fact (as many emergentists and epiphenomenalists claim), or their existence in the natural world must be denied altogether (as substance dualists and eliminative physicalists claim). If there is hylomorphic structure, however, the options are no longer constrained in the foregoing way. Distinctive powers like yours and mine exist in the natural world because structure exists in the natural world. Moreover, because structure is a basic principle on the hylomorphic view, this does not simply push the demand for an explanation back a step. A framework's basic principles stand in need of no further explanation within that framework. Structure and things that get structured are both basic on the hylomorphic view. Nothing must explain why the former exists any more than something must explain why the latter does. As a result, the view leaves it unmysterious why and how mental phenomena exist in the natural world.

Despite differing in its basic principles from familiar mind-body theories, hylomorphism has a familiar profile. On the one hand, it is naturalistic; it claims that we are physical beings with physical components, and that our distinctive powers to think, feel, perceive, and act are essentially embodied in the physical materials that compose us. On the other hand, it is antireductive; it denies that descriptions and explanations of those powers are reducible to the descriptions and explanations provided by physics, chemistry, or neuroscience. The reason is that there is more to having and exercising those powers than the materials and conditions that lower-level sciences describe and explain. In particular, there are the higher-level structures that unify simpler lower-level entities and occurrences into more complex individuals and activities. These structures are what delineate the subject matters of special sciences such as biology and psychology, and what secure the autonomy those sciences enjoy.

Articulating an acceptable mind-body theory that is both antireductive and broadly naturalistic has been a major objective of philosophy of mind for the past fifty years. Familiar theories of this sort include various forms of nonreductive physicalism and emergentism. The kind of hylomorphic theory I defend is antireductive and naturalistic as well, but it rejects both physicalism and some of the central tenets of emergentism. It thus makes a unique contribution to the ongoing effort to articulate an acceptable nonreductive naturalism. It is nevertheless similar enough in its outlines to these more familiar theories that their exponents are likely to see in hylomorphism a congenial alternative or a worthy competitor. Hylomorphism also has unique resources for solving the problems that beleaguer competing theories; these include the problem of emergence, the problem of downward causation, and the problem of other minds. I've described hylomorphic solutions to mind-body problems elsewhere (Jaworski 2011: chapters 10–11; 2012), but I've not described in detail the metaphysics underlying those solutions. This book aims to do precisely that.

Hylomorphic metaphysics has been much neglected since the seventeenth century. As a result, the biggest challenge facing hylomorphists is to articulate their view with enough detail and rigor to enable it to compete with more familiar alternatives. Often when philosophers discuss hylomorphism they approach it historically in terms of what Aristotle, Aquinas, Leibniz, or some other philosopher of the past has claimed. That is not my approach. The hylomorphic theory I defend dovetails with current work in metaphysics, philosophy of mind, and scientific disciplines such as biology and neuroscience. I argue that there are good philosophical and empirical reasons to reintroduce hylomorphism's core notion of organization or structure. The descriptions and explanations of living behavior advanced by many biologists, neuroscientists, and other empirical practitioners posit notions of organization or structure that play theoretical roles like the hylomorphic notion. If we are committed to countenancing the entities postulated by our best descriptions and explanations of reality, and we think those descriptions and explanations derive from empirical sources such as the sciences, then scientific appeals to structure make a serious ontological demand. The most direct way of meeting that demand is to take scientific appeals to structure at face value—to claim that structure really exists, and that it plays the theoretical roles those appeals imply. We can express those roles with some slogans:

Structure matters: it operates as an irreducible ontological principle, one that accounts at least in part for what things essentially are.

Structure makes a difference: it operates as an irreducible explanatory principle, one that accounts at least in part for what things can do, the powers they have.

Structure counts: it explains the unity of composite things, including the persistence of one and the same living individual through the dynamic influx and efflux of matter and energy that characterize many of its interactions with the wider world.

I articulate and defend a hylomorphic theory built around a notion of structure that plays these roles. That theory is based on a substance-attribute ontology which takes properties to be powers and tropes. The result differs substantially from the hylomorphic theories defended recently by Kit Fine (1999; 2008), Mark Johnston (2006), Kathryn Koslicki (2008), David Oderberg (2007), Michael Rea (2011), and Anna Marmodoro (2013). It is nevertheless similar in its implications to the hylomorphic theory defended by Robert Koons (2014) and to the theory Montgomery Furth (1978: 638–9) attributes to Aristotle.

I argue that structures are powers to organize or configure things—powers that structured wholes are essentially engaged in manifesting. You and I are essentially engaged in configuring the materials that compose us; we impose a human-wise organization on them, and we persist exactly as long as we do so. These particular configurings—yours and mine—are particularized properties or tropes. They are numerically different properties that nevertheless resemble each other rather closely—more closely than, say, either resembles Fido's configuring the materials that compose him or the oak tree's configuring the materials that compose it. The hylomorphic view I defend weds this account of structure to an account of composition similar in its outlines to Peter van Inwagen's (1990). For van Inwagen, composition happens exactly when the activities of fundamental physical particles constitute a life. Importantly, lives on van Inwagen's view play many of the theoretical roles that hylomorphic structures do: *lives matter* for van Inwagen; *lives make a difference*, and *lives count*. It is thus easy to see the hylomorphic view of composition as an extension of van Inwagen's. According to hylomorphists, composition happens exactly when there is an individual that configures materials, and the configuring activities in which individuals engage are like lives in van Inwagen's sense.

The hylomorphic view nevertheless differs from van Inwagen's in three respects. First, it is not committed to atomism as van Inwagen's is, but remains neutral on the nature of fundamental physical materials. Second, it is open to the possibility that there might be structured individuals other than living things, something van Inwagen denies. Third, unlike van Inwagen, hylomorphism endorses the existence of functional parts such as eyes, hearts, and brains. Van Inwagen claims that the literal details about what lives are and what characteristics they have need to be supplied empirically (1990: 84). Contrary to his view of composition, however, the actual empirical work of biologists, neuroscientists, and others indicate that there are parts such as eyes, hearts, and brains. Central to that work is the method of functional analysis, which looks to analyze the activities of complex wholes into simpler subactivities performed by simpler subsystems. Parts emerge as hierarchically arranged systems and subsystems that are identified and individuated by the contributions they make to the activities of whole organisms. They are, in other words, subpockets of order within more inclusive ordered wholes. The foregoing differences between van Inwagen's view of composition and the hylomorphic one enable the latter, I argue, to respond more effectively to the objections that have been levied against the former.

One important point of convergence with van Inwagen is that composite individuals are *emergent* on the hylomorphic view: they have distinctive powers different from the powers of unstructured materials. It is this feature that enables hylomorphists to solve the mind-body problems that beleaguer their competitors. The key to understanding the hylomorphic approach to mind-body problems is the notion of an *activity-making structure*. The structures mentioned above, the ones that are like van Inwagen's lives, are structures that make individuals what they are; they are *individual-making structures*, the kinds of things traditional hylomorphists called 'substantial forms.' But individual-making structures are not the only structures that exist on the hylomorphic view. The activities in which structured individuals engage have structures as well; they are activity-making structures. The idea that there are activity-making structures is based on the observation that the activities of structured individuals involve coordinated manifestations of the powers of their parts. When we walk, talk, sing, dance, reach, grasp, run, jump, throw, breathe, and engage in the various other activities we do, we are imposing an order on the ways our parts manifest their powers. On the hylomorphic view, these structured manifestations of powers include thinking, feeling, and perceiving. These activities, like the ones just mentioned, are essentially embodied in the physiological mechanisms that compose us, yet it is not possible to reduce descriptions and explanations of them to descriptions and explanations of physiological mechanisms. The reason, again, is that there is more to these activities on the hylomorphic view than the operations of physiological mechanisms; there is also the way those operations are coordinated or structured, and structure in general is something different from things that are structured. The hylomorphic view is thus robustly antireductive despite its commitment to essential physical embodiment.

There are several philosophical challenges facing the hylomorphic theory I articulate. Some of them concern the metaphysics of powers and tropes on which it is based. I defend the power-trope metaphysic in Chapters 2–5 against alternatives like David Armstrong's, which posit categorical properties and universals. I do not claim that the metaphysics I defend is the only one capable of supporting a hylomorphic theory. I do not even claim that it is the best. I claim only that it provides a workable basis for understanding hylomorphism and its implications for the philosophy of mind.

Other challenges to the hylomorphic view concern its account of composition. They include the following:

(1) Arguments such as those advanced by Peter van Inwagen (1981) and Trenton Merricks (2001a), which purport to show that there are no functional parts such as eyes, hearts, and brains.

(2) Arguments such as those advanced by Ted Sider (1993) and Dean Zimmerman (2003), which purport to show that a view of composition similar to van Inwagen's has problems accommodating the possibility of gunk (infinitely

divisible stuff), and also problems accommodating a commonsense ontology that includes artifacts and natural bodies such as mountains and planets.

(3) Arguments such as those advanced by David Lewis (1986b) and Ted Sider (2001), which purport to show that composition cannot be restricted.

(4) Arguments advanced by Thomists such as David Oderberg (2007) and Anna Marmodoro (2013), which purport to show that things like electrons cannot survive being incorporated into more inclusive wholes, that they are instead obliterated and replaced by proxies which are essentially biological.

I defend the hylomorphic view of composition against these arguments in Chapters 6 and 7.

In Chapter 8, I defend the *embodiment thesis*, the claim that the capacities of structured wholes are essentially embodied in their parts. The embodiment thesis, I argue, represents the default hylomorphic position. It is challenged by Aristotelian and Thomistic hylomorphists who claim that the operation of one capacity in particular, thought or understanding (Aristotle's *nous*, Aquinas' *intellectus*), is not essentially embodied. The traditional arguments for this claim derive from Aristotle's argument in *De Anima* III.4. I argue in line with other commentators that the argument is flawed in multiple ways.

Another challenge derives from the worry that hylomorphism is, as Bernard Williams once put it, "just a polite form of materialism" (1986: 224), where 'materialism' is synonymous with 'physicalism.' There are several versions of Williams' worry. Some focus on hylomorphism's commitment to supervenience; others on its commitment to exhaustive physical decomposition, and yet others on its approach to mental phenomena—phenomenal consciousness in particular. I argue that each version of Williams' worry is based on a false premise. Several versions of Williams' worry derive from confusions about physicalism and reductionism, as well as supervenience, necessitation, and their relationship to explanation. I argue that many definitions of physicalism are based on dubious metaphysical assumptions. These include definitions of nonreductive physicalism that appeal to property dualism (Kim 2006a: 13), and definitions of physicalism that appeal to supervenience or necessitation (Chalmers 1996; Stoljar 2010). These definitions, I argue, fail to capture the core physicalist thesis, the claim that everything is physical. As a result, they fail to express accurately what physicalism is. Something analogous is true of various definitions of reduction. They fail to express accurately what reduction requires. In addition, many philosophers wrongly assume that necessitation and supervenience imply some type of explanatory condition; that if A supervenes upon or is necessitated by B, then B somehow provides an explanation for A, and that this explanation undermines the causal or explanatory autonomy of A. The assumption is false for reasons that are well rehearsed in the literature. I clarify the foregoing notions in Chapters 9–11, and show that as a result, several versions of Williams' worry disappear.

A related challenge comes from philosophers like David Chalmers (2009, 2010). Chalmers argues that any theory which approaches phenomenal consciousness in the way hylomorphism does must be false. Hylomorphists nevertheless reject a key premise in Chalmers' argument, namely the premise that conceivability is in general a guide to possibility. There are many circumstances in which we form conceptions of things that do not correspond to any possible situations. This approach to Chalmers' argument requires what he calls 'strong necessities'—a posteriori truths that have metaphysically necessary primary intensions. Chalmers advances several arguments against strong necessities. Their crux is the premise that every conceivable scenario corresponds to a possible world. Hylomorphists reject this premise. According to them, there are many circumstances in which we form conceptions that correspond to no possible situations. I discuss Williams' worry and Chalmers' challenge in Chapters 12 and 13.

In Chapters 13 and 14, I argue against competing versions of hylomorphism such as those endorsed by Kit Fine, Mark Johnston, and Kathryn Koslicki, and argue that the version of hylomorphism I defend is a worthy competitor in metaphysics and the philosophy of mind. Its theoretical merits are at least as choiceworthy as those of its closest rivals: nonreductive physicalism, emergentism, and Russellian monism. Part of what makes hylomorphism attractive is that it implies elegant solutions to mind-body problems such as the problem of emergence, the problem of mental causation, and the problem of other minds. Other philosophers have looked to advance similar solutions, but many have fallen short because they have failed to situate their solutions within a broader metaphysical framework. Tyler Burge (1993, 2006) is an example. He has looked to solve the problem of mental causation by appeal to commonsense explanatory practice, but because he has not situated that practice within a metaphysical framework that explains why it should be what it is, his attempt to solve the problem has failed to address the deeper metaphysical puzzles about mental causation (Kim 1998). Hylomorphism provides the needed framework.

Hylomorphism is also able to make sense of many of our intuitions about composition, about the persistence of organisms, and about the distinctive powers they possess. Moreover, it is easier to defend hylomorphism than it is to defend physicalism because it does not make the totalizing claim about physics or any other branch of science that physicalism does. Because hylomorphism endorses a plurality of structures in the natural world, it endorses a plurality of autonomous scientific disciplines for describing them. Consequently, it does not insist that everything can be exhaustively described and explained by a single conceptual framework the way monistic theories like physicalism do. These points, combined with hylomorphism's ability to solve a number of persistent philosophical problems, make a strong case in its favor. At the very least, they show that hylomorphism deserves a place at the table alongside the more familiar alternatives.

1

Structure in the World

1.1 The Hylomorphic Notion of Structure

This book is about hylomorphism, and hylomorphism is about structure. Hylomorphism claims that structure (or organization, form, arrangement, order, or configuration) is a basic ontological and explanatory principle. Some individuals, paradigmatically living things, consist of materials that are structured or organized in various ways. You and I are not mere quantities of physical materials; we are quantities of physical materials with a certain organization or structure. That structure is responsible for us being and persisting as humans (as opposed to, say, dogs or rocks), and it is responsible for us having the particular developmental, metabolic, reproductive, perceptive, and cognitive capacities we have.

The hylomorphic notion of structure is not the same as others that have appeared in the literature. It is not the same, for instance, as the notion of structure that has been operative in discussions of grounding in metaphysics (Schaffer 2009; Sider 2012). 'Structure' in that sense refers to what grounds the distinction between things that are fundamental or perfectly natural in Lewis' (1983b) sense, and things that are not. Intuitively, we want to say that a predicate like 'is green' corresponds more closely to what there really is—to the world's structure—than a predicate like 'is grue.' Hylomorphism provides one account of what structure in the grounding sense includes, but the specifically hylomorphic notion of structure (or organization, form, arrangement, order, or configuration) is not the same as that notion. Nor is the hylomorphic notion of structure the same as the notion that is operative in debates about scientific realism (Worrall 1989; Ladyman and Ross 2007). 'Structure' in those debates refers to the relational contents of scientific theories that remain constant across episodes of theory change. But that is not what 'structure' in the hylomorphic sense refers to. Nor is the hylomorphic notion of structure the same as the one David Chalmers sometimes employs:

> Physical descriptions of the world characterize the world in terms of structure and dynamics . . . A microphysical description of the world specifies a distribution of particles, fields, and waves in space and time. These basic systems are characterized by their spatiotemporal properties, and properties such as mass, charge, and quantum wavefunction state. These latter properties are ultimately defined in terms of spaces of states that have a certain abstract structure (e.g., the space of continuously varying real quantities, or of Hilbert space states), such that the states play

a certain causal role with respect to other states. We can subsume spatiotemporal descriptions and descriptions in terms of properties in these formal spaces under the rubric of *structural* descriptions. The state of the systems can change over time in accord with dynamic principles defined over the relevant properties. The result is a description of the world in terms of its underlying spatiotemporal and formal structure, and dynamic evolution over the structure…Truths about consciousness are not truths about structure and dynamics. (Chalmers 2002a: 258)

For Chalmers, structural descriptions are contrasted with dynamic ones; roughly, structural descriptions use exclusively logical, mathematical, and spatial terms to specify a system's state at a time, whereas dynamic descriptions add temporal and nomic terms to specify how a system's states change over time. For hylomorphists, structure is neither this general nor this abstract. Nor are structural descriptions limited to the conceptual resources of physics. On the contrary, for hylomorphists, paradigmatic structural descriptions derive from biology.

The hylomorphic notion of structure is closer to the notion of organization that Armstrong sometimes employs:

For a Materialist, a man is a physical object, distinguished from other physical objects only by the special complexity of his physical organization. (Armstrong 1968: 11)

There is nevertheless a key difference between Armstrong's notion of organization or structure and the hylomorphic one. For Armstrong, the kinds of organization that characterize living things like us can be countenanced unproblematically by physics; there is no sense that the existence of biological organization might pose a challenge to the physicalist idea that everything can be exhaustively described and explained by physics. Hylomorphists, as we will see, disagree.

To help illustrate the hylomorphic notion of structure, I'll use three examples. The first is *the squashing example*. Suppose we put Godehard in a strong bag—a very strong bag since we want to ensure that nothing leaks out when we squash him with several tons of force. Before the squashing, the contents of the bag include one human being; after, they include none. In addition, before the squashing the contents of the bag can think, feel, and act, but after the squashing they can't. What explains these differences in the contents of the bag pre-squashing and post-squashing? The physical materials (whether particles or stuffs) remain the same—none of them leaked out. Intuitively, we want to say that what changed was the way those materials were structured or organized. That organization or structure was responsible for there being a human before the squashing, and for that human having the capacities it had. Once that structure was destroyed, there no longer was a human with those capacities. Structure is thus a basic ontological principle: it concerns what things there are. It is also a basic explanatory principle: it concerns what things can do.

A second example introduces hylomorphism by contrasting it with a more familiar view: physicalism. Physicalism claims that everything is physical: everything can be

exhaustively described and explained in principle by physics. Philosophers sometimes use the term 'physicalism' to refer to much weaker claims, such as the claim that everything has physical properties, that all properties supervene on physical properties, or that everything is composed of physical parts. These weaker claims (which I'll discuss in detail in Chapter 11) are compatible with things having nonphysical properties in addition to physical ones. Physicalism, in the strong sense I have in mind, rules this out. According to physicalism in the strong sense, every individual, every feature it has, and every behavior in which it engages can be described and explained exhaustively using only the conceptual resources of physics. It is possible, in other words, at least in principle, for physics to provide us with descriptions and explanations that leave out nothing but give us the complete story about what there is, what it does, how it does it, and why.

To appreciate more fully what physicalism implies, imagine a character: the super physicist, a being that possesses complete physical knowledge of the universe. The super physicist has complete knowledge of all the fundamental physical entities in the universe: what they are, what properties they have, what relations they stand in, and what laws govern their behavior. Imagine, however, that the super physicist lacks a psychological conceptual framework and even a biological one. It lacks the perceptual and conceptual resources to distinguish living things from nonliving ones, or mental beings from nonmental ones. The concepts of life, perception, desire, belief, money, sex, and so on are completely beyond its ken. As a result, when it describes the universe, its descriptions are framed solely in the vocabulary of physics—solely in terms of the characteristics of fundamental physical particles or stuffs. Because the super physicist does not have the concepts to distinguish living things from nonliving ones, or mental beings from nonmental ones, its descriptions make no mention of plants, animals, or people, nor do they mention any distinctive biological or psychological activities such as growth, reproduction, perception, or belief. Nor can the super physicist recognize the distinctions these things mark in the natural world. It recognizes no difference between Godehard and the surrounding air, for instance; from its standpoint, there is just a continuous curtain of fundamental physical particles or stuffs.

Many people would be inclined to say that the super physicist's descriptions of the world would be missing some very important things: the distinction between life and nonlife or between intelligence and nonintelligence, not to mention the stock of things that tend to occupy most people's minds, such as money, food, sex, family, health, and professional success. If physicalism is true, however, the super physicist's description misses nothing. Since everything can be exhaustively described and explained by physics, the super physicist's descriptions of the universe are complete as they stand. If you and I describe the universe in ways that recognize the distinctions between living things and nonliving ones, or mental beings and nonmental ones, that is a comment not necessarily about what the universe contains, but about how we go about describing it.

Hylomorphists disagree. They claim that the super physicist's descriptions of the world are in fact missing some very important things, namely the various ways that physical materials are structured or organized in the natural world—ways that mark the difference between living things and nonliving ones, or between mental beings and nonmental ones, that distinguish Godehard from the surrounding air, that confer on the particles or stuffs located exactly where he is the unity that makes him a distinctive whole, and that enable him to persist one and the same over time.

Someone might object that my description of physicalism is a straw man, that the physicalist can accommodate the kinds of natural differences that I've claimed are distinctive of hylomorphism. Part of what I'll argue in later chapters is that this objection is mistaken. Once physicalism is formulated more precisely, it cannot accommodate the kinds of natural differences hylomorphism can.

A third way of illustrating the basic hylomorphic notion of structure involves the many appeals to structure we find in the sciences, especially biology and biological subdisciplines such as neuroscience. Scientists frequently appeal to notions of structure, order, arrangement, or configuration, and at least some of these appeals appear to be ontologically serious; that is, they appear to posit structure as a real ontological and explanatory principle. Here is an example taken from a popular college-level biology textbook—note the references to organization, order, arrangement, and related things:

> Life is highly organized into a hierarchy of structural levels, with each level building on the levels below it . . . Biological order exists at all levels . . . [A]toms . . . are ordered into complex biological molecules . . . the molecules of life are arranged into minute structures called organelles, which are in turn the components of cells. Cells are [in turn] subunits of organisms . . . The organism we recognize as an animal or plant is not a random collection of individual cells, but a multicellular cooperative . . . Identifying biological organization at its many levels is fundamental to the study of life . . . With each step upward in the hierarchy of biological order, novel properties emerge that were not present at the simpler levels of organization . . . A molecule such as a protein has attributes not exhibited by any of its component atoms, and a cell is certainly much more than a bag of molecules. If the intricate organization of the human brain is disrupted by a head injury, that organ will cease to function properly . . . And an organism is a living whole greater than the sum of its parts . . . [W]e cannot fully explain a higher level of order by breaking it down into its parts. (Campbell 1996: 2–4)

This passage suggests that the way things are structured, organized, or arranged plays an important role in them being the kinds of things they are, and in explaining the kinds of things they can do. There could not be an organism without the right 'cooperative' arrangement or organization of cells. Moreover, what the organism and structured wholes like it can do depends on that organization: destroy the organization of the brain, and you destroy its functional capacities. Organization, order, structure, or arrangement is thus something real, with real explanatory significance.

This idea is echoed by others. Consider William Bechtel, a philosopher of neuroscience:

[T]he organization of ... components typically integrates them into an entity that has an identity of its own ... Organization itself is not something inherent in the parts ... Accordingly, investigators who already understand in detail how the parts behave are often surprised by what happens when they are organized in particular ways ... In virtue of being organized systems, mechanisms do things beyond what their components do ... Not only can one study the performance of a mechanism without knowing its component parts and their operations, but what the mechanism as a whole does is typically quite different than the operations performed by its parts ... As a result, organized mechanisms become the focus of relatively autonomous disciplines ... This autonomy maintains that psychology and other special sciences study phenomena that are outside the scope of more basic sciences but which determine the conditions under which lower-level components interact. In contrast, the lower-level inquiries focus on how the components of mechanisms operate when in those conditions ... The fact that mechanisms perform different activities than do their parts manifests itself in the fact that the activities of whole mechanisms are typically described in different vocabulary [sic] than are component operations. Traditional accounts of theory reduction implicitly recognized this fact by requiring bridge principles to connect the different vocabularies used in different sciences, but little notice was given as to why different sciences employ different vocabularies. The vocabulary used in each science describes different types of entities and different operations—one describes the parts and what they do, whereas another describes the whole system and what it does. (2007: 174, 185–6)

According to Bechtel, a complex whole—what he calls a 'mechanism'—consists of parts plus an organization that confers on it capacities not had by its parts taken in isolation. In addition, Bechtel brings out a related point: descriptions of structured wholes and explanations of their behavior are irreducible to descriptions and explanations of their unstructured parts. This was implicit in Campbell's claims that complex wholes have novel properties not seen at lower levels of organization, and that we do not fully explain the behavior of complex wholes by breaking them down into their parts. Bechtel makes the point explicit. If he is right, then higher-level empirical disciplines and lower-level ones have different subject matters on account of the ways things are organized or structured. Moreover, because higher- and lower-level disciplines deal with different subject matters, they have different vocabularies, and provide different kinds of explanations, something that makes higher-level disciplines autonomous—irreducible to lower-level disciplines in the traditional philosophical sense.

Nancy Cartwright makes similar points about structure (though her preferred terms are 'arrangement' and 'configuration'). For her, the notion of structure is closely connected to the notion of a nomological machine:

Law-like regularities result from the successful operation of nomological machines (77) ... What is a nomological machine? It is a fixed ... arrangement of components, or factors, with stable ... capacities that in the right sort of stable ... environment will, with repeated

operation, give rise to the kind of regular behavior that we represent in our scientific laws (49–50) . . . Sometimes God supplies the arrangements . . . but very often we must supply them ourselves. (1999: 122)

Moreover, Cartwright (1999: 128–30) argues for an antireductive thesis similar to Bechtel's. Attempts to describe and explain the behavior of nomological machines in terms of causal chains alone, she says, fail to recognize that appeals to the arrangement of something's components and appeals to events that stand in causal chains answer different kinds of questions and aim to explain different kinds of things. Appeals to structure aim at explaining the stable behavior of the machine; appeals to causal chains, on the other hand, aim to explain more localized antecedent triggers for some event or other. As a result, appeals to causal chains cannot take over the descriptive and explanatory jobs that appeals to arrangements or configurations perform.

Bechtel, Campbell, and Cartwright are not alone in conceiving of organization or structure as a real and irreducible ontological and explanatory principle. Consider the biologists Camazine et al. and their distinguished forebears, J. B. S. Haldane, G. G. Simpson, and Ernst Mayr, respectively:

[P]attern formation often is achieved by systems without external guidance. The mechanism of self-organization in biological systems differs from those in physical systems in two basic ways. The first is the greater complexity of the subunits in biological systems . . . The second difference concerns the nature of the rules governing interactions among system components. In chemical and physical systems, pattern is created through interactions based solely on physical laws . . . Of course, biological systems obey the laws of physics, but in addition . . . the subunits in biological systems acquire information about the local properties of the system and *behave* according to particular genetic programs that have been subjected to natural selection. This adds an extra dimension to self-organization . . . because in [biological] systems selection can finely tune the rules of interaction. By tuning the rules, selection shapes the patterns that are formed and thus the products of group activity can be adaptive . . . [I]nteractions among system components can be surprisingly simple, even when extremely sophisticated patterns are built. (Camazine et al. 2001: 12–13)

[L]ife is essentially a pattern of chemical happenings . . . What is common to all life is the chemical events . . . [T]he chemical changes which go on in the leaves, bark, and roots of a tree . . . are surprisingly like those which go on in human organs . . . In fact, all life is characterized by a fundamentally similar set of chemical processes arranged in very different patterns . . . Shakespeare's plays consist of words . . . It is important to know this, as it is important to know that life consists of chemical processes. But the arrangement of the words is even more important than the words themselves. And in the same way life is a pattern of chemical processes . . . [E]nzymes and other proteins can be purified and will carry on their characteristic activities in glass bottles. And no biochemist would say they were alive . . . [T]o suppose that one can describe life fully on [chemical] lines is to attempt to reduce it to a mechanism, which I believe to be impossible. On the other hand, to say that life does not consist of chemical processes is to my mind as futile and untrue as to say that poetry does not consist of words. (Haldane 1947: 54–6)

[T]o understand organisms one must explain their organization . . . [O]ne must know what is organized and how it is organized . . . The aim of biology is to understand the structure, functioning . . . and history of organisms. (Simpson 1964: 113)

All biologists are thorough-going "materialists" in the sense that they recognize no supernatural or immaterial forces, but only such that are physico-chemical . . . [T]he modern biologist rejects in any form whatsoever the notion that a "vital force" exists in living organisms which does not obey the laws of physics and chemistry. All processes in organisms, from the interaction of molecules to the complex functions of the brain and other whole organs, strictly obey these physical laws . . . But [modern biologists] do not accept the naïve mechanistic explanation of the seventeenth century and disagree with the statement that animals are "nothing but" machines . . . Where organisms differ from inanimate matter is in the organization of their systems. Organismic biologists stress the fact that organisms have many characteristics that are without parallel in the world of inanimate objects. The explanatory equipment of the physical sciences is insufficient to explain complex living systems. (Mayr 1982: 2, 52)

Mayr makes explicit a further implication of structure being an ontological and explanatory principle: structure is the basis for at least some of the divisions we recognize within the natural world—the division, for instance, between living things and nonliving ones. Consider likewise the cyberneticist Gerd Sommerhoff and the philosopher of biology Michael Ruse; both look to ground natural distinctions in the organization or structure things have:

The physico-chemical picture of the living organism is only half the truth. The missing half concerns the nature of the organizational relationships that make the behavior of obviously living systems uniquely different from that of obviously non-living systems . . . In many ways this is the more important half. For here lie the differences between life and death, and between higher and lower forms of life as they affect us most . . . Even if we knew down to the last molecular detail what goes on inside a living organism, we should still be up against the fact that a living system is an organized whole which by virtue of the distinctive nature of its organization shows unique forms of behaviour which must be studied and understood at their own level. (Sommerhoff 1969: 147–8)

The whole point about the DNA molecule—the point at which Watson and Crick hinted at the end of their classic paper—is that you get every different kind of gene . . . not from different subunits, but from the same subunits ordered in the chain in different ways. That is what the genetic "code" is all about. The same is true of the brain. Order the molecules in one way and you get junk. Order the molecules in another way and you get William Shakespeare . . . The order exists, it is not unreal, but it is not a thing in the way that a molecule is a thing. To think otherwise is to get oneself into that way of thinking which gives existence to such very nonuseful entities as . . . *élans vitaux* . . . The assembled and functioning DNA molecule is not a new substance. It is smaller substances (or substance parts), ordered . . . And analogous comments apply to us humans. The very crux of the Darwinian explanation of the distinctiveness of humankind is that we are ordered, and thus can function in ways that are not possible for other animals. (Ruse 2001: 79)

When people think of structure, they often think of something static, such as the relatively unchanging spatial relations among atoms in a crystal. In fact, even some

contemporary hylomorphists use the term ‘structure’ in this way. David Oderberg is an example:

> [M]ere structure in the sense of configuration of parts is far too *static* a concept to tell you all there is about the form of an animal: There are its characteristic functions and behaviour, its dispositions, instincts, tendencies, actions and reactions, and all the rest of which ethology is made. These *dynamic* notions have to be added to the relatively static structural notions to get us to something like an account of the form of a living thing. (2014: 177)

But the philosophers and scientists we’re considering don’t view structure so narrowly. Although we can refer to the sum of spatial relations among something’s parts as a structure, the structures that are likely to interest us most—the kinds of structures that, say, distinguish living things from nonliving ones—are not static spatial relations, but dynamic patterns of environmental interaction, the sorts of things that Oderberg refers to as ‘forms.’ The neurophysiologist Jonathan Miller brings out the idea of dynamic structure in the following way:

> [T]he physical universe tends towards a state of uniform disorder ... In such a world the survival of form depends on ... [either] the intrinsic stability of the materials from which the object is made, or the energetic replenishment and reorganisation of the material which is constantly flowing through it ... The configuration of a fountain ... is intrinsically unstable, and it can retain its shape only by endlessly renewing the material which constitutes it; that is, by organising and imposing structure on the unremitting flow of its own substance ... The persistence of a living organism is an achievement of the same order as that of a fountain ... it can maintain its configuration only by flowing through a system which is capable of reorganising and renewing the configuration from one moment to the next. But the engine which keeps a fountain aloft exists independently of the watery form for which it is responsible, whereas the engine which supports and maintains the form of a living organism is an inherent part of its characteristic structure. (Miller 1978: 140–1)

The biologist J. Z. Young brings out the idea that distinctively living structures are dynamic by describing living organization as a complex activity:

> The essence of a living thing is that it consists of atoms ... caught up into the living system and made part of it for a while. The living activity takes them up and organizes them in its characteristic way. The life of a man consists essentially in the activity he imposes upon that stuff. (Young 1971: 86–7)

Similarly, Peter van Inwagen describes the dynamic character of living structure using the metaphor of a self-maintaining storm:

> The events and processes that collectively constitute an organism’s being alive might be described in [the following] terms by some disembodied intellect that knew its physics and chemistry but which had never heard of organic life and which was observing its first living organism:
>
> What I am observing is an unimaginably complex self-maintaining storm of atoms ... I would compare it with the Great Red Spot on Jupiter, which has been in existence for hundreds

of years. (Or I might compare it with a wave, or the propagation of a wave, which is a sort of self-maintaining event that involves different particles of fluid at different times.) This storm moves across the surface of the world, drawing swirls and clots of atoms into it and expelling others, always maintaining its overall structure. One might call it a homeodynamic event. This observation is an acute one. There are events such as these. They are what I call lives. (van Inwagen 1990: 86–7)

Mark Johnston calls the kinds of structures that Miller, Young, and van Inwagen have in mind "dynamic principles of unity," ones that comprise programmatic sequences of changes over time, and that often involve different kinds of changes under different kinds of conditions:

A form or principle of unity may be . . . *dynamic*, in that . . . the parts it holds of vary over time . . . A paradigm case is a living thing whose organic matter is unified into an organism by some categorical basis of a multi-track disposition to such life-function as ingestion, assimilation, excretion, growth, metabolic repair, and so on and so forth. In this case, the principle of unity is a complex structure of biochemical relations, whose holding of the organic matter of the living thing provides the categorical basis for the multi-track disposition in question. The operation of that disposition *requires* the matter to be exchanged over time. (2006: 663–4)[1]

Johnston, Miller, and van Inwagen bring out an additional point. The dynamic structure that qualifies something as a living thing is also what enables that thing to persist through time. It is one and the same organism that persists through the constant influx and efflux of matter and energy because of its structure and its dynamic ability to impose that structure on incoming matter and energy. John Locke made a similar point:

[C]onsider wherein an oak differs from a mass of matter . . . [O]ne is only the cohesion of particles of matter any how united, the other such a disposition of them as constitutes the parts of an oak; and such an organization of those parts as is fit to receive and distribute nourishment, so as to continue and frame the wood, bark, and leaves, *etc*., of an oak, in which consists the vegetable life . . . For this organization . . . *is* that individual life . . . The case is not so much different in *brutes* . . . Something we have like this in machines . . . [F]or example, what is a watch? . . . [N]othing but a fit organization or construction of parts to a certain end . . . If we would suppose this machine one continued body, all whose organized parts were repaired, increased, or diminished by a constant addition or separation of insensible parts, with one common life, we should have something very much like the body of an animal . . . An animal is a living organized body; and consequently the same animal . . . is the same continued life communicated to different particles of matter, as they happened successively to be united to that organized living body . . . This also shows wherein the identity of the same man consists; viz. in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body. (Locke 1959: Book II, chapter 27, sections 5–9)

[1] The theory of structures I defend nevertheless differs in significant ways from Fine's and Johnston's theories, which I discuss in more detail in Chapter 14.

A final idea about structure is introduced by John Dewey. He suggests that a view of the natural world that recognizes organization or structure as an irreducible ontological and explanatory principle might have important implications for understanding mental phenomena:

> The difference between the animate plant and the inanimate iron molecule is not that the former has something in addition to the physico-chemical energy; it lies in the *way* in which physico-chemical energies are interconnected and operate . . . Iron as a genuine constituent of an organized body acts so as to tend to maintain the type of activity of the organism to which it belongs. If we identify . . . the physical as such with the inanimate we need another word to denote the activity of organisms as such. Psycho-physical is an appropriate term . . . In the compound word, the prefix 'psycho' denotes that physical activity has acquired additional properties . . . Psycho-physical does not denote an abrogation of the physico-chemical; nor a peculiar mixture of something physical and something psychical . . . it denotes the possession of certain qualities and efficacies not displayed by the inanimate. Thus conceived there is no problem of the relation of physical *and* psychic. There are specifiable empirical events marked by distinctive qualities and efficacies. There is first of all, *organization* . . . Each 'part' of an organism is itself organized, and so of the 'parts' of the part . . . 'mind' is an added property assumed by a feeling creature, when it reaches that organized interaction with other living creatures which is language, communication. (Dewey 1958: 253–8)

Dewey suggests the possibility of understanding mental phenomena as species of structural phenomena in general. Suppose that he and the other authors we've considered are right, and that structure is uncontroversially part of the natural world. If mental phenomena are just species of structural phenomena, they must be uncontroversially part of the natural world as well. The account of mental phenomena Dewey suggests thus appears to be naturalistic. Moreover, because structural descriptions and explanations are irreducible on this view to descriptions and explanations that appeal to unstructured physical materials, the account also appears to be antireductive. Finally, if structure is a basic ingredient of the natural world as Dewey and the other authors suggest, then it is unmysterious on this account how mental phenomena could be both natural and irreducible.

I take the foregoing passages to gesture toward a certain view of structure in the natural world. According to it, structure is something real with ontological and explanatory significance. It serves as a principle of unity, persistence, and power. It is responsible for setting something apart as a discrete individual distinct from the rest of the physical universe. It explains why such an individual can exist one and the same over time. It also explains why that individual can do many of the things it does: why it has many of the powers it has, including the powers that classify it as a living being or a mental one. Finally, an individual's structure explains the autonomy of various empirical disciplines that would look to describe and explain its behavior. We can express the theoretical roles that structure plays on this view with some slogans:

Structure matters: it operates as an irreducible ontological principle, one that accounts at least in part for what things essentially are.

Structure makes a difference: it operates as an irreducible explanatory principle, one that accounts at least in part for what things can do, the powers they have.

Structure counts: it explains the unity of composite things, including the persistence of one and the same living individual through the dynamic influx and efflux of matter and energy that characterize many of its interactions with the wider world.

Structure minds: it provides us with resources for understanding the place of mental phenomena within the natural world.

In the chapters that follow, I plan to develop this view of structure in greater detail. I'll distinguish the central notion of structure from other notions of structure, such as the broader notion employed by Ted Sider and the narrower ones employed by David Armstrong and David Chalmers. I'll argue, perhaps surprisingly, that the view I describe is incompatible with physicalism. I'll nevertheless argue that it is a view many would-be physicalists would likely find congenial; it in fact closely resembles some varieties of nonreductive physicalism. Many would-be emergentists would also likely find it congenial since it closely resembles some classic forms of emergentism. I will nevertheless argue that the view has advantages over both nonreductive physicalism and emergentism since it gingerly sidesteps many of the problems they have faced, including the problem of emergence and the problem of downward causation. I'll argue finally that we have good empirical and philosophical reasons to think that the view, or something very much like it, is true. Even if you disagree with this conclusion in the end, I hope to develop the view in enough detail that it will be clear exactly where the disagreement lies, and this by itself is a kind of progress.

1.2 Four Approaches to Structure

Many philosophers find claims about organization or structure like those discussed in Section 1.1 obvious and unremarkable: How could anyone question (and hence why would anyone bother to mention) that structure factors into things? But the notion of structure does not come for free, at least not if we endorse *ontological naturalism*, the idea that when it comes to determining what exists, empirical investigation—paradigmatically science—is our best guide. John Dupré has endorsed a similar thesis:

> I place myself firmly in the philosophical tradition that sees empirical, often scientific, inquiry as providing the most credible source of knowledge of how things are. (Dupré 1993: 1)

Replacing the phrase "how things are" in Dupré's statement with "what there is" yields what I am calling 'ontological naturalism.'

Ontological naturalism can be understood as the conjunction of a broadly Quinean thesis about ontological commitment with a broad empiricism. The broadly Quinean thesis maintains that we are committed to all the entities postulated by our best descriptions and explanations of reality, and a broad empiricism maintains that

our best descriptions and explanations of reality derive from empirical sources such as the natural and social sciences. Suppose we take the natural-language sentences in which our best descriptions and explanations are formulated and reformulate those sentences in a quantifier-variable idiom the way Quine (1948) suggests. In that case, says the Quinean thesis, we would be committed to the existence of all the entities needed to make those descriptions and explanations true. Two remarks are in order.

First, ontological naturalism does not imply that the sciences are the only sources of empirical knowledge. It does not rule out a guiding role for commonsense or even legend or myth in determining what exists, nor does it imply that empirical knowledge is our only guide to what exists. It is weaker than Wilfrid Sellars' (1963a: 173) *scientia mensura*: science is the measure of all things, of what is that it is and of what is not that it is not. Ontological naturalism does nevertheless accord to empirical sources a privileged role in determining what exists, and it takes the sciences as paradigmatic examples of such sources.

Second, it's likely that our understanding of ontological commitment should be expanded somewhat beyond Quine's original boundaries. Empirical methods and techniques appear to provide important sources of ontological commitment as well.[2] The idea is roughly that in constructing and executing experiments we are often implicitly committed to a range of assumptions that carry ontological commitments of their own but that might nevertheless remain unstated in the more canonical descriptions and explanations we give. If this is the case, then we need to expand the basic Quinean thesis to accommodate these further commitments. On this expanded understanding, ontological naturalism says that we are committed not only to the existence of the entities needed to make our best empirical descriptions and explanations true, but also to the existence of the entities needed to make our best empirical methods and techniques effective.

I plan to take ontological naturalism as a working assumption in what follows. Defending it would take us beyond the scope of the present inquiry, so I will leave its defense for a later date. Its significance for our purposes is this: If our best empirical descriptions, explanations, and methods posit entities of kind *K*, then that gives us good prima facie reason to think that *K*s exist. Consider, then, empirical descriptions and explanations like those discussed earlier—descriptions and explanations that posit various kinds of structure, order, arrangement, organization, or configuration. If ontological naturalism is true, these descriptions and explanations give us good prima facie reason to think those structures exist. They therefore make of us a serious ontological demand. The most straightforward way of meeting this demand takes empirical claims about structure at face value. It says that structure really is an

[2] The idea that the nature of things is revealed not just in description and explanation, but in the methods or techniques we use to study them, is defended by Hacking (1983), and it plays a major role in many of Cartwright's (1999) arguments.

irreducible ontological and explanatory principle. This straightforward realist approach is the one favored by hylomorphists. But there are at least three alternatives to it that are worth mentioning. They have in common the assumption that everything can be exhaustively described and explained *without* appeal to structure.

First, *structure eliminativism* claims that empirical statements about structures like the ones discussed earlier are literally all false. In reality, say eliminativists, there is no such thing as structure or organization, nor is talk of structure or organization useful for serious descriptive and explanatory purposes—the sorts of purposes that drive scientific endeavor. Talk of structure or organization is instead the byproduct of a defective way of trying to describe and explain the behavior of things. When we finally give a complete scientific account of what living or mental beings are and why they behave as they do, that account will not appeal to structure or organization any more than a complete scientific account of the weather appeals to the Greek gods.

Second, *structure reductivism* claims that at least some empirical statements about structure are true, and that appeals to structure can do serious descriptive and explanatory work, but according to structure reductivists, statements about structure can do this only because the structures they postulate can be identified with conditions that can be exhaustively described and explained without appeal to structure. The structure of the human brain, for instance, can be identified in principle with some nonstructural condition *C*, and as a result, say reductivists, claims about human brain structure can be rewritten as claims about condition *C* by the substitution rule for identity. This is true across the board for structural discourse, say reductivists; structures can in general be identified with nonstructural conditions. It is thus possible in principle for nonstructural discourse to take over all the descriptive and explanatory jobs that appeals to structure perform. Structural discourse is thus reducible to nonstructural discourse.

A third view, *nonreductive structure antirealism*, tries to steer a middle course between reductivism and eliminativism. Like reductivists and unlike eliminativists, nonreductive structure antirealists try to countenance talk of structure, but like eliminativists and unlike reductivists, they try to avoid countenancing the entities to which structural discourse appears to commit us. Nonreductive structure antirealists agree with reductivists that structural discourse has genuine descriptive and explanatory legitimacy, but they reject the idea that this legitimacy is grounded in the identification of structures with nonstructural conditions. It is grounded instead in our descriptive and explanatory interests—interests that nonstructural discourse is incapable of satisfying. Even though everything can be exhaustively described and explained in nonstructural terms, say nonreductive structure antirealists, we often have descriptive and explanatory interests that cannot be satisfied unless we employ talk of structure. Appeals to structure are useful, therefore, not because there is something other than what can be exhaustively described and explained without appeal to structure, nor because structures can be identified with nonstructural

conditions; appeals to structure are useful rather because they satisfy descriptive and explanatory interests that would otherwise go unsatisfied.

Nonreductive structure antirealism is the kind of view Mark Johnston describes in the following terms and subsequently rejects:

> When certain items come to stand in certain relations . . . there then comes to be some further item which has those original items as parts. That is . . . how we have such complex items as model airplanes, trains, and molecules . . . [J]ust why are those relations . . . "item-generators," while other relations . . . seem impotent in the production of new items? . . . Could it just be a projection of our idiosyncratic way of experiencing and conceptualizing reality, so that things considered in themselves are not complex, but are so only relative to a scheme of clumping or bundling? Somehow, I doubt it. (2006: 652)

But perhaps the best example of nonreductive structure antirealism is the view Daniel Dennett (1991) calls 'mild realism' about patterns. According to Dennett, we postulate patterns with an eye to predicting the behavior of things in ways that are more efficient (if less precise) than the conceptual resources of physics allow. Patterns can be considered real to the extent that postulating them allows us to make predictions more efficiently (if less precisely) than physics. According to Dennett, however, it is possible for two different and incompatible conceptual frameworks to postulate patterns, and yet for there to be no way of determining which framework is more accurate. "The choice of a pattern," he says, "would . . . be up to the observer, a matter to be decided on idiosyncratic pragmatic grounds" (1991: 49). For Dennett, then, whether there are patterns and what patterns there are is ultimately determined by our practical idiosyncrasies.

Nonreductive structure antirealists look to retain the sparse nonstructural ontology that motivates eliminativism and reductivism without either denying the legitimacy of structure talk or carrying the empirical burden of identifying structures with nonstructural things. According to them, our descriptive and explanatory interests ground the legitimacy of structural discourse without positing entities other than those posited by nonstructural discourse, and without issuing promissory notes for the identification of structural entities with nonstructural ones. Instead, they promise to provide paraphrases of statements about structure that aim at minimizing their ontological seriousness. When, for instance, neuroscientists say that the debilitating effects of head injuries result from disruptions to the delicate structure of the brain, nonreductive structure antirealists look to describe and explain the debilitating effects of the head injury without appealing to the structure of the brain. Neuroscientists employ talk of brain structure, they say, simply because it satisfies certain special interests that neuroscientists have—interests in, say, formulating counterfactual-supporting generalizations that cannot be formulated in nonstructural terms (Fodor 1974: 114), or in predicting behavior in a way that is more efficient than predicting it without appeal to structures (Dennett 1991: 37ff.). Whatever the details, nonreductive structure antirealists will look to countenance the neuroscientist's talk of structure—its legitimacy,

usefulness, or success—without compromising their commitment to the idea that everything can be exhaustively described and explained without appeal to structure, and without identifying structures with nonstructural conditions.

Importantly, nonreductive structure antirealism implies that structural discourse is not reducible to nonstructural discourse. Even if the world is fundamentally nonstructural, say nonreductive structure antirealists, we still describe it in many different ways that satisfy many different kinds of interests, including interests different from those that nonstructural discourse satisfies. If nonstructural discourse cannot satisfy all the descriptive and explanatory interests we have—if, in particular, it cannot satisfy the interests we use structural discourse to satisfy—then nonstructural discourse is incapable of taking over the descriptive and explanatory roles that structural discourse plays.[3] Structural discourse, in other words, is irreducible to nonstructural discourse, not because there is something other than what can be exhaustively described and explained in nonstructural terms, but rather because we have special interests that the conceptual resources of nonstructural discourse cannot satisfy.

Let us now return to the very first approach to structure mentioned above—the straightforward realist approach that takes empirical claims about structure at face value. Unlike the approaches just described, it claims that appeals to structure cannot be reduced to, or paraphrased, or eliminated in favor of nonstructural descriptions and explanations. Call this view *structure realism*.

Structure realism should not be confused with *structural* realism, which is a position in debates about scientific realism (Ladyman and Ross 2007). Structural realism claims that scientific theories have structures that remain invariant across instances of theory change. If would-be scientific realists commit themselves only to those structures, say structural realists, and not to the ontologies of particular theories, then realism will be immune to antirealist arguments that appeal to theory change. This view is altogether different from what I am calling *structure* realism. Structure realism, in the sense I intend, is not a position in the scientific realism debate. It is instead defined by its rejection of the assumption the aforementioned three approaches to structure have in common, namely the assumption that everything can be exhaustively described and explained without structure.

Structure realists agree with nonreductive structure antirealists that structural discourse is irreducible to nonstructural discourse, but they anchor their antireductionism in metaphysical soil. Structure, they say, is a real and irreducible ontological and

[3] Another argument to this effect appeals to intertheoretic identities: If theory T_A is reducible to theory T_B, then T_B can take over all the descriptive and explanatory jobs of T_A. But this kind of takeover requires that entities postulated by T_A be identical to entities postulated by T_B (Sklar 1967; Schaffner 1967; Causey 1977: chapter 4; Jaworski 2011: 123–4). According to nonreductive structure antirealists, however, structural and nonstructural discourse are not related in straightforward ways that would allow us to identify structures with nonstructural things. Consequently, nonstructural discourse cannot take over the descriptive and explanatory roles of structural discourse, and hence structural discourse is not reducible to nonstructural discourse. For more details on the notion of reduction, see Section 11.3.

explanatory principle that exists independent of our descriptive and explanatory interests, and that is distinct from anything nonstructural. Like structure reductivists, moreover, structure realists take the ontological implications of structural discourse at face value; they do not think it is possible to paraphrase appeals to structure in ways that minimize the seriousness of those implications. Unlike reductivists, however, they deny that structures can be identified with things that can be exhaustively described and explained in nonstructural terms. Structure realists thus countenance talk of structure together with the entities to which that talk appears to commit us, but they do so by denying that everything can be exhaustively described and explained without appeal to structure. Structure, they say, is a real feature of things distinct from any nonstructural things that exist. Descriptions and explanations that appeal to structure cannot be reduced to claims about nonstructural conditions, nor can they be accurately paraphrased in ways that do not appeal to structure.

1.3 Structure Realism and Hylomorphism

Structure realism by itself is compatible with a wide variety of positions on what there is. It is, for instance, compatible both with physicalism and with substance dualism. Consider the former. Physicalism claims that everything is physical; everything can be exhaustively described and explained by our best current or future physics. Suppose that physicalism is true, and that our best physics ends up postulating real, irreducible structures. In that case, structure realism ends up being true as well. Structure realism is thus compatible with physicalism. Consider now substance dualism. It claims that persons, such as you and I, are not physical entities; we are not, for instance, human organisms. Suppose that substance dualism is true, and suppose further that human organisms consist of physical materials that are structured in certain ways. Structure, in other words, is a basic principle that factors into descriptions of what human organisms are and what they can do; it simply doesn't factor into descriptions of what we, nonphysical persons, are and what we can do. The upshot is a substance dualistic view that is committed to structure realism. Structure realism is thus compatible with substance dualism.

Yet, neither physicalism nor substance dualism accurately represents the kind of view expressed in the foregoing passages from biologists. Contrary to substance dualism, that view is committed to the claim that we are organisms or at least physical beings of some sort. Contrary to physicalism, on the other hand, that view is committed to the claim that there are basic structures other than those postulated by physics. It distinguishes what physics by itself can describe and explain from what appeals to biological, psychological, or social structure enable us to describe and explain. Because of this, its exponents claim that organisms are not mere machines, as Mayr puts it, but have characteristics—"emergent properties," to use Campbell's term—not found among nonliving things. This leads them to conclude in turn that the explanatory apparatus of physics is insufficient to describe and explain living

behavior, that physics gets at "only half the truth," as Sommerhoff puts it. What physics misses, say exponents of this view, are things that can only be described and explained by appeal to structure at a biological, psychological, or social level. I'll reserve the term 'hylomorphism' for a structure realist view of this sort.

'Hylomorphism' is a compound of the Greek words *hyle* and *morphe*, which are typically translated as 'matter' and 'form', respectively. Because the label is not new, it's worth mentioning that the hylomorphic view I have in mind differs from those of Kit Fine (1999; 2008), Mark Johnston (2006), David Oderberg (2007), Kathryn Koslicki (2008), Michael Rea (2011), Anna Marmodoro (2013), and Robert Koons (2014). Exactly how it differs and how it is similar is something that will emerge in the course of the subsequent discussion. In addition, it is worth mentioning that I cannot vouch for the view's similarities to those of Aristotle, Aquinas, Leibniz, Merleau-Ponty, or the other philosophers whose views have been labeled 'hylomorphic.'[4] The view is nevertheless very similar to the one Montgomery Furth attributes to Aristotle:

> [The world] is an Empedoclean, finite three-dimensional mass, entirely filled with the four elements . . . [W]e now observe that scattered through this three-dimensional mass there are innumerable *knots*, regions where the matter is elaborately and intricately worked up into an organic unity . . . highly convoluted but relatively stable eddies in the general commingling-and-separation [of elements] . . . Aristotle thinks the "principle" called "form" must be brought in on top of the Empedoclean basis, to explain the stability of the knots and the complex specific character that they manifest as long as they last . . . a material individual (i.e., animal) is . . . a semipermanent warp or bend informing the local matter, which the matter flows through at various rates during the organism's life history . . . while the form imposes the continuity. (1978: 638–9)

One difference between Aristotle's view and the hylomorphic view I have in mind is that the latter is not committed to a specific account of fundamental physical entities. It does not claim that they are Empedoclean stuffs, as Aristotle does, but is happy to leave it to physicists to determine what they are. I'll have more to say on this point in Chapter 6. For the time being, however, the similarities between these views are more important than their differences. According to both, fundamental physical materials of some sort (whether Empedoclean stuffs, subatomic particles, or something else entirely) get structured into distinctive organic beings—relatively stable 'knots' or 'eddies,' as Furth puts it, in the overall flow of fundamental physical material. Again, it is helpful to understand this view by contrasting it with physicalism.

Suppose the physical universe is a vast sea of matter and energy that is described or will be described by our best physics. Physicalism claims that this vast sea is all there

[4] For Aristotle's view, see *Physics*, Book II, chapters 1–3, and *On the Soul*, as well as Nussbaum and Putnam 1992. For Aquinas' view, see *Summa Theologiae* Ia, questions 75–86, and Leftow 2001; for more on Leibniz, see Garber 1985 and Smith 2002, and see Merleau-Ponty (2002) for his view.

is. When hylomorphists look at the world, they see the same sea of matter and energy that physicalists do, but they see something more besides: scattered throughout it are tiny localized pockets of order or arrangement—semistable, self-maintaining patterns of flow, like eddies in a river, with distinctive structures or forms. According to hylomorphists, the behavior of these structured pockets or eddies cannot be described and explained using the conceptual resources of physics alone—not even in principle. Physics enables us to describe the matter and energy that flow through structured individuals. It enables us to describe the interactions among the fundamental physical particles or stuffs in them, but it does not enable us to describe the overall pattern that makes this collection of fundamental physical interactions a human, a dog, or an oak tree. Understanding the physical materials that are caught up into these pockets or eddies will tell us something about the characteristics and behavior of the whole, but it won't tell us everything. It won't tell us, for instance, what sets the whole apart as a unified individual distinct from the inanimate materials that surround it. The materials, after all, can exist without being caught up into the whole, and the whole can exist without the materials that are caught up into it at any given moment. In addition, the materials by themselves won't tell us why the whole possesses powers that the materials by themselves do not possess, such as the powers to perceive, or think, or feel. There must be some further principle that accounts for the unity and also the persistence of the whole that explains why numerically one and the same whole can exist at different times despite the constant flux of materials that are caught up into it, and that explains the distinctive powers of living things. That principle, say hylomorphists, is structure, the persistent dynamic structures we find in living things. These structures are what confer unity on the various material bits that are caught up into them, and they are what confer on living wholes powers not had by the material bits taken in isolation. Those structures, hylomorphists insist, can only be described using the conceptual resources of biology. And depending on the kinds of living things we are talking about, something analogous might be true of other special sciences such as psychology and economics. Describing the patterns of environmental interaction that set perception, feeling, thought, goal-directed movement, and the other activities of complex living things apart from the rest of the ebb and flow of matter and energy in the universe might require the distinctive conceptual resources these other sciences provide.

Let these remarks serve as a first rough sketch of the kind of view hylomorphism represents. My goal in the chapters that follow is to describe the view in greater detail. I'll begin by articulating a basic metaphysical framework in Chapters 2–5. I use that framework to develop an account of structure in Chapters 6–10, and to define physicalism and other mind-body theories in Chapter 11. In Chapter 12, I distinguish physicalism from hylomorphism and address what I call Williams' worry: the suspicion once expressed by Bernard Williams (1986) that hylomorphism is just a 'polite' form of physicalism. I argue that this is not the case. I also explain why

hylomorphism is not committed to reductivism, despite being committed to robust necessitation and supervenience theses. In Chapter 13, I explain how a general hylomorphic framework can be used to address topics in the philosophy of mind, and how it solves the problem of emergence, the problem of downward causation, and the problem of other minds. Finally, in Chapter 14, I make the case that hylomorphism deserves to be taken at least as seriously as views like nonreductive physicalism, emergentism, and Russellian monism.

2

Individuals, Properties, and Events

2.1 A Basic Substance-Attribute Ontology

The goal of this chapter is to introduce a basic metaphysical framework for the discussions that follow. I think that many confusions in philosophy of mind have resulted from philosophers failing to make their metaphysical assumptions explicit. I hope to make my metaphysical assumptions as transparent as possible. I'll begin by sketching a basic ontology. Its outlines are described by David Armstrong:

> [T]he world consists simply of particulars having properties and relations[;] it is particulars and particulars alone which can act and be acted upon. But . . . they act and are acted on solely in virtue of their properties, non-relational and relational . . . Another way of putting this is that it is states of affairs which are causes. (Armstrong 1978a: 132)

What Armstrong describes here is a substance-attribute ontology, one that takes substances or *individuals* as I'll call them (he calls them 'particulars'), and attributes or *properties*, to be the fundamental entities.[1] I use the term 'property' broadly here to encompass not merely monadic properties, but *n*-adic properties (or relations) generally. Having a mass of 30.1kg is a property expressed by the monadic predicate 'has a mass of 30.1kg,' and being 3cm taller than is a property expressed by the relational predicate 'is 3cm taller than.'

A substance-attribute ontology stands opposed to ontologies that would aim to reduce one of these kinds of entities to the other. Ontologies that aim to reduce individuals to properties include bundle theories of substance (Stout 1921, 1923; Williams 1953, 1986; Campbell 1981, 1983, 1990; Simons 1994; Mumford 2012). Ontologies that aim to reduce properties to individuals include extreme forms of nominalism (Goodman and Quine 1947; Quine 1953; Goodman 1956; Sellars 1963b). Unlike both, a substance-attribute ontology takes individuals and properties to be irreducibly distinct and equally fundamental.

Armstrong mentions a third ontological category in addition to individuals and properties, what he calls 'states of affairs,' instances of individuals having properties

[1] Some philosophers use the term 'substance-attribute ontology' to refer to an ontology committed to universals. Mumford (2012) is an example. The ontology I've outlined in this chapter is compatible with properties being universals, but it is not committed to this.

(or standing in relations). Some philosophers call these entities 'facts' (Papineau 1993; Chalmers 1996; Carruthers 2000; Bird 2007), but I will typically follow philosophers who call them 'events' (Goldman 1970; Kim 1973, 1976; Bennett 1988). I will, however, occasionally help myself to Armstrong's term and call them 'states of affairs.' Unlike individuals and properties, states of affairs or events are not fundamental entities since they depend on entities of the other sorts; they consist of individuals having properties or standing in relations. An event exists exactly if an individual has a property at a time or several individuals stand in a relation at a time. Event e is identical to event e^* exactly if e and e^* comprise the same individuals, properties, and times. If, for instance, event e is identical to a's having property P at time t, and event e^* is identical to b's having property Q at time t^*, then $e = e^*$ exactly if $a = b$, $P = Q$, and $t = t^*$.[2]

Several notes are in order about this view of events. First, it does not stand opposed, except perhaps terminologically, to views that reserve the term 'event' for changes, cases in which an individual gains or loses a property (Lombard 1986, 1998). These cases can easily be accommodated as a subset of events in the foregoing sense, and if need be, we can use the term 'changes' to refer to them.[3] This view of events does nevertheless stand opposed to several others. First, it stands opposed to views that take events to be fundamental entities. Davidson (1969) suggests a view like this; though he does not commit himself to it (1970: 210). Second, it stands opposed to views that take events to be bundles of properties. Ted Honderich (1982, 1988) appears to endorse a view of events like this, and Thomas Nagel (1986) sometimes suggests one. Third, it stands opposed to views like Chisholm's (1970), which take events to be universals instead of unrepeatable particulars. Fourth, it stands opposed to views that imply conditions for event identity and individuation different from those just described. Davidson (1969), for instance, individuates events not by their constitutive individuals, properties, and times, but by their causes and effects: event e_1 is identical to event e_2, on Davidson's view, if and only if e_1 and e_2 have exactly the same causes and effects. Likewise, even though Fodor (1974: 100) suggests that events consist of individuals having properties at times, the theoretical work he expects events to perform implies that property identity is not necessary for event identity as it is on this account (how event_1 can be identical

[2] Some philosophers reserve the term 'event' to refer to changes. Lombard (1986) is an example. Clearly, not all events in our sense involve changes: Alexander's being seventy inches tall right now is an event in our sense, but does not involve a change. If it becomes necessary, I propose to use the term 'changes' to refer just to those events that involve changes.

[3] Lombard essentially concedes this terminological point: "My own view of events should, I suppose, be numbered among the property exemplification accounts . . . An event is . . . the *exemplifying* of a *dynamic* property" (1998: 289); an event, in other words, is the exemplifying of a specific kind of property, a dynamic one. Lombard draws a distinction between being an exemplification and being an exemplifying in order to distinguish his view of events from those of Kim and Bennett. This distinction collapses, however, with an ontology of tropes, something I'll discuss momentarily.

to event$_2$ without having exactly the same properties is something he doesn't endeavor to explain).[4]

Finally, I will argue later that events needn't be posited as a separate ontological category as Armstrong (1989a: 88–9; 1997: 115ff.) would have it. If properties are tropes, as I'll argue in Chapter 3, then individuals having properties is sufficient to do all the truthmaking work that Armstrong posits states of affairs or events to perform (Section 3.2).

2.2 The Eleatic Principle

The principal agents in this ontology are individuals. They can both act on other individuals and be acted on by them. Individuals, in other words, have powers that can manifest themselves both *actively* and *passively*—both in the ways they affect things and in the ways they are affected by them. Individuals enter into causal relations by exercising their powers, and they are empowered in the ways they are by their properties.

Properties play several theoretical roles in this ontology. First, they are causal enablers: they confer the powers that make causal interactions among individuals possible. Second, because properties confer powers, they are also causal explainers: they explain why individuals act or are acted on in the ways they are. Third, properties ground the objective similarities and differences among individuals. Individuals are always similar or dissimilar in certain respects. These respects are properties (Martin 1996a: 71–3; 1997; 2007: 42–3). Moreover, since properties are possessed by individuals independent of any descriptive or explanatory interests we happen to have, the similarities and differences they ground are objective.

This view of individuals and properties fits naturally with an epistemology (Campbell 1981, 1990; Martin 1997, 2007): we know about individuals by the ways they interact with each other, that is, the ways they manifest or exercise their powers. These include especially the ways in which they exercise their powers on us, and in which we exercise our powers on them. We know about the properties of individuals in the same way. One and the same individual has the power to affect

[4] Fodor says: "Token physicalism is simply the claim that all events that the sciences talk about are physical events . . . [T]oken physicalism is weaker than what might be called 'type physicalism,' the doctrine, roughly, that every *property* mentioned in the laws of any science is a physical property. If every event is the instantiation of a property, then type physicalism does entail token physicalism; two events will be identical [sic] when they consist of the instantiation of the same property by the same individual at the same time . . . [Yet] Token physicalism does not entail type physicalism . . . because the contingent identity of a pair [sic] of events presumably does not guarantee the identity of the properties whose instantiation constitutes the events" (1974: 100). Putting aside Fodor's suggestion that two things could be identical, which I take to be a slip of the tongue, Fodor suggests that events are property instantiations, yet his argument that token physicalism does not entail type physicalism succeeds only if event identity does not require property identity. He must therefore be committed to an identity condition for events other than the one described above.

us in different ways because it is red and also round. It has the powers to affect us in different ways on account of having different properties: redness and roundness. What exactly the natures of these properties are is an open empirical question, one that we might feel confident answering only after we have investigated them a good deal further.

Based on what I've said so far, it should be evident that the properties I have in mind are *natural* properties, not mathematical or logical ones. For our purposes, we can put properties of the latter sorts to one side. When it comes to natural properties and the individuals having them, I follow Armstrong (1978b: 46) and others in taking a cue from the Eleatic stranger in Plato's *Sophist*:

> I suggest that anything has real being, that is so constituted as to possess any sort of power either to affect anything else or to be affected . . . I am proposing as a mark to distinguish real things, that they are nothing but power. (247d–e)

I'll follow Graham Oddie (1982) in calling this the *Eleatic Principle* (Kim (1993b: 348) calls it 'Alexander's dictum'). Roughly, the Eleatic principle says that the only things that exist are ones that play some sort of causal role. When the Eleatic principle is applied to a substance-attribute ontology, it implies that the only individuals that exist are ones that can enter into causal relations, and the only properties that exist are ones that empower individuals to enter into those relations. This claim implies that properties are *sparse* as opposed to *abundant*.

The terms 'sparse' and 'abundant' were introduced by David Lewis (1983b). Because Lewis took properties to be abundant, he didn't speak of sparse properties but of universals in Armstrong's sense (1978a, b). What he said about Armstrong's universals could nevertheless have applied equally well to tropes as understood by philosophers such as George Molnar (2003), John Heil (2003), and C. B. Martin (2007):

> [U]niversals are sparse. There are the universals that there must be to ground the objective resemblances and the causal powers of things . . . The guiding idea . . . is that the world's universals should comprise a minimal basis for characterizing the world completely . . . A satisfactory inventory of universals is a non-linguistic counterpart of a primitive vocabulary for a language capable of describing the world exhaustively . . . It is quite otherwise with properties. Any class of things, be it ever so gerrymandered and miscellaneous and indescribable in thought and language, and be it ever so superfluous in characterizing the world, is nevertheless a property . . . Because properties are so abundant, they are undiscriminating . . . [P]roperties do nothing to capture facts of resemblance . . . Likewise, properties do nothing to capture the causal powers of things. Almost all properties are causally irrelevant. (Lewis 1983b: 12–13)

In this book, I will *not* follow Lewis' use of the term 'property.' I will use the term 'property' instead to refer to sparse properties. Properties in this sense are what explain the objective similarities and differences among individuals and the causal powers individuals have. Given the Eleatic principle, moreover, there are only those

properties whose existence would be sufficient to account for those similarities, differences, and powers. If it becomes necessary to refer to what Lewis calls 'properties' in his abundant sense, I will use the term 'sets' or 'classes,' or sometimes just 'abundant properties.'

The claim that properties are sparse has several noteworthy implications. First, properties do not correspond to predicates one-one. Predicates can be identified with sentence-frames (Strawson 1974: 37–8; Armstrong 1978a: 2–3), linguistic expressions such as '___ is red' and '___ is taller than ___' that form sentences when the blanks are filled in by terms. (For convenience, I'll omit the blanks when talking about predicates in the future.) Properties are supposed to be the nonlinguistic correlates of at least some predicates. According to some accounts of abundant properties, they are the nonlinguistic correlates of every predicate; every predicate expresses a property, and different predicates express different properties (Shoemaker (1980) calls properties in this sense 'Cambridge properties'). This is not the case if properties are sparse.

If properties are sparse, then it is possible for different predicates to express the same property ('weighs 453.59 grams,' 'weighs 1 pound'), and for different properties to be expressed by the same predicate in different contexts ('The team is good,' 'The wine is good') (Armstrong 1978b: 9–14; Campbell 1990: 25; Molnar 2003: 26). Likewise, if properties are sparse, not every predicate expresses a property. For one thing, there might be unknown properties to which no actual predicates correspond (Armstrong 1978b: 12–14; Molnar 2003: 25). In addition, some predicates are self-referentially incoherent, such as the predicate 'is a property to which no predicate corresponds,' which generates a version of Russell's paradox. If properties are sparse, these predicates do not (and indeed cannot) correspond to any property (Molnar 2003: 26). It is also possible to invent predicates, but if properties are sparse, it is not possible to invent properties. When we invent predicates, moreover, we can do so by iterating formal operations ad infinitum, yet it is implausible to suppose that the number of properties could be infinite (Ellis and Lierse 1994: 9; Molnar 2003: 26). Similarly, we can invent a nondenumerably infinite number of predicates to describe physical entities such as subatomic particles, provided we take seriously the use of real numbers in physics. But it is extremely implausible to suppose that those physical entities should have nondenumerably infinite properties (Bradley 1979: 12–13; Molnar 2003: 26). For all of these reasons, sparse properties do not correspond one-one to predicates.

If properties are sparse, therefore, not all predicates express properties; moreover, different predicates can express the same property, and different properties can be expressed by the same predicate in different contexts. What, then, determines which predicates express sparse properties? What, moreover, determines which properties those predicates express, and whether they are the same or different properties that are expressed by other predicates? According to the account of properties I've been outlining, the answers to these questions derive largely from empirical sources

(cf. Swoyer 1982: 205; Ellis 2002: 44–5; Molnar 2003: 27; Armstrong 2010: 19).[5] Recall the thesis of ontological naturalism (Section 1.2). It implies that determining what properties exist and which predicates express which properties is largely an empirical undertaking. We take our best empirical descriptions, explanations, methods, and techniques, and countenance all the properties needed to make the descriptions and explanations true and the methods and techniques effective.

In the sections that follow, I will argue that there are good reasons to prefer sparse properties to abundant ones when addressing issues in the philosophy of mind.

2.3 An Argument against Abundant Properties

Theories that endorse abundant property appear to be of broadly two sorts. The first sort takes properties to be identical to or at least coextensive with predicates. The second takes properties to be classes. I'll follow Armstrong (1978a, b; 1989a) in referring to these as 'predicate nominalism' and 'class nominalism,' respectively. One argument against abundant properties claims that predicate and class nominalism are both unacceptable, and as a result, we should reject abundant properties in favor of sparse ones.

Predicate nominalism has been endorsed by many philosophers. According to Prior et al., for instance, "*every* expression like 'the property of being *F*' denotes a property" (1982: 254, emphasis in original). One difficulty with predicate nominalism is that some predicates are self-referentially incoherent, such as the predicate 'is both red all over and not red all over' (Molnar 2003: 26). Predicate nominalists seem committed to these predicates expressing properties, yet surely something cannot be both red all over and not red all over. Likewise, it seems possible that there might be properties to which no actual predicates correspond (Armstrong 1978a: 17; 1978b: 12–14; 1989a: 11; Molnar 2003: 25). If, for instance, the predicate 'is white' had never been invented, it seems that white objects could still exist.

Other problems with predicate nominalism concern the identity conditions of properties. Either different predicates express different properties according to the view or not. Suppose the former is true. In that case, predicate nominalism claims that the predicates 'has a mass of 1kg' and 'has a mass of 1,000 grams' express different properties. But that seems absurd; surely a kilogram is identical to 1,000 grams. Suppose, then, that different predicates can express the same property. In that case,

[5] Brian Ellis states the basic idea as follows: "[W]hich predicates designate properties[?] . . . I answer: first decide what properties and structures you must postulate if you wish to give an adequate account of the phenomena, and then decide which expressions of the language refer to these properties or structures" (Ellis 2002: 44–5). Consider likewise Armstrong, Molnar, and Swoyer, respectively: "[H]ow do we determine what the true universals are? My suggestion is that they are best postulated on the basis of *total science*" (Armstrong 2010: 19); "[W]hat properties there are is . . . determined . . . on a posteriori grounds, most likely by current best science" (Molnar 2003: 27); "the claim that there are such things as properties is a philosophical one, but determining just what properties there are is—like questions about existence generally—an empirical matter" (Swoyer 1982: 205).

the predicate nominalist must provide further criteria for determining when different predicates express the same property and when they do not. Yet, it is difficult to see how they could provide these criteria without abandoning their theory in favor of another. Suppose, for example, that they claim that the predicate '*P*' and the predicate '*Q*' express the same property exactly if '*P*' and '*Q*' apply to exactly the same individuals; '*P*' and '*Q*,' in other words, are coextensive. In that case, they appear to have abandoned predicate nominalism in favor of class nominalism. Likewise, suppose that they claim that predicate '*P*' and predicate '*Q*' express the same property exactly if the individuals in the extensions of '*P*' and '*Q*' have exactly the same causal powers. In that case, they appear to have abandoned abundant properties in favor of sparse ones. Analogous worries attend the possibility that the same predicate might express different properties as when we say 'The team is good' and 'The wine is good' (Armstrong 1978b: 9–14; Campbell 1990: 25; Molnar 2003: 26).

In addition, as we saw a moment ago, we can invent predicates. It seems counterintuitive, however, to suppose that we can invent properties. Among other reasons, we are capable of inventing an infinite number of predicates simply by iterating formal operations, but it seems implausible that there could be an infinite number of properties (Bradley 1979: 12–13; Ellis and Lierse 1994: 9; Molnar 2003: 26). These are some of the reasons to reject predicate nominalism.

Class nominalism has been defended most forcefully by David Lewis (1983b, 1986a), as well as by philosophers who take properties to be intensions, that is, functions from possible worlds to extensions (Montague 1960: 152; Egan 2004). There are two well-rehearsed arguments against class nominalism. The first is the coextension problem (Goodman 1966; Armstrong 1989a: 25–6; Campbell 1990: 33; Loux 2006: 76). If class nominalism is true, then properties are simply classes of particulars. Different properties will therefore correspond to different classes of particulars, and vice versa. The problem is that it seems that different properties can correspond to the same class of particulars. For example, the class of creatures with a heart is coextensive with the class of creatures with a kidney. If class nominalism is true, and properties are classes, this would seem to imply that the property of having a heart is identical to the property of having a kidney, but surely this is false, and in that case, class nominalism must be false.

The second problem with class nominalism is the bi-extension problem (Armstrong 1978a: 46ff.; Armstrong 1989a: 27; Wolterstorff 1960; 1970: chapter 8). The coextension problem claims that two different properties can have the same extension; the bi-extension problem claims that the same property can have different extensions. Consider the class of all actual red objects. It includes this apple, that stop sign, the tomatoes in the garden, and so on. According to class nominalists, this class is identical to the property of being red. Suppose, however, that one of these objects had never existed. In that case, it seems to follow that, according to class nominalism, the property of being red would not have existed either. But that seems absurd. Intuitively, we want to say that the property of being red would still exist; it would just

have had a different extension. If the actual extension of red corresponds to the class {*a*, *b*, *c*}, and in counterfactual world *w* object *a* does not exist, it surely does not follow that red does not exist in *w*. Surely *b* and *c* would still be red; surely their redness would not be affected by the absence of some other red object. Class nominalism implies this counterintuitive result, critics say, because it identifies a property with the class of objects that are in its extension.

There is a way class nominalists can respond to these arguments. It appeals to David Lewis' modal realism. According to Lewis, modal claims such as 'Wittgenstein could have been an aeronaut' are made true by counterparts (Lewis 1968; 1971; 1973a: 39–43; 1986a: chapter 4). Counterparts, according to Lewis, are individuals who exist at nonactual possible worlds and who resemble actual individuals more closely than any other individuals at those worlds. The properties and behavior of counterparts are what make modal claims about actual individuals true. According to Lewis, for instance, Wittgenstein has a counterpart (call him 'Wittgenstein*') who exists at some possible world *w*. Wittgenstein* resembles Wittgenstein more closely than any other object at *w* does, and he is also an aeronaut. This makes it true of the actual Wittgenstein that he could have been an aeronaut, for he has a counterpart at some possible world who is an aeronaut. The rub of Lewis' view is that in order for counterparts to play this truthmaking role they must exist. On Lewis' view, Wittgenstein* the aeronaut is just as real as Wittgenstein the philosopher; he simply doesn't exist at the actual world the way Wittgenstein does. Many philosophers consider this kind of extreme modal realism highly implausible, but it does provide class nominalists with a solution to the coextension and bi-extension problems.

Those problems are based on the ideas that it is possible for different properties to have the same extension, and for the same property to have different extensions. Lewis' modal realism allows class nominalists to rule out these possibilities. Because there are nonactual individuals on Lewis' view, nominalists can define properties by their extensions not merely in the actual world, but across all possible worlds. Even though it might be true that all actual creatures with a heart are creatures with a kidney, it is still possible that there could be a creature with a heart that does not have a kidney. If that is the case, then there is at least one possible world in which some creature belongs to the extension of 'creature with a heart' but not to the extension of 'creature with a kidney.' Suppose, then, that class nominalists define the property of being a creature with a heart as the class of all possible individuals (not merely actual ones) that have hearts; and suppose, likewise, that they define the property of being a creature with a kidney as the class of all possible individuals (not merely actual ones) that have a kidney. Since these classes are not necessarily coextensive, since there is at least one possible world in which not every creature with a heart is a creature with a kidney, class nominalists can countenance the distinction between being a creature with a heart and being a creature with a kidney, for even if these properties have the same extension in the actual world, they have different extensions across all possible worlds.

This might seem like good news for class nominalism, but many philosophers would consider this a Pyrrhic victory. The reason is that most reject Lewis' extreme modal realism. Consequently, the foregoing response is not an option for them. Second, when it comes to the philosophy of mind in particular, most philosophers will not want the claims they make in that domain to depend on metaphysical assumptions as extreme and contentious as Lewis'. The foregoing response is not an option for these philosophers either. On the whole, then, there appear to be good reasons to reject both predicate nominalism and class nominalism, especially when it comes to the philosophy of mind. But since abundant property theories are typically committed to one of these views, it seems that there are good reasons to reject abundant properties as well, especially when it comes to the philosophy of mind.

The foregoing argument against abundant properties is by no means knockdown. But together with the independent arguments for sparse properties advanced in the next section, it provides some grounds for rejecting abundant properties in favor of sparse ones, especially when it comes to the philosophy of mind.

2.4 Arguments in Favor of Sparse Properties

There are several arguments in favor of sparse properties. Alexander Bird (2007: 10–12) discusses four of them. First, our intuitions tend to favor sparse properties, something evinced by the sense that something like being grue is not a genuine property but a fabrication. Although intuitions are not indefeasible, they do have some evidential value. The reason, Bird argues, is that they have a basis in natural selection:

> Our cognitive systems . . . have evolved in large part as property-detection systems: animals have evolved capacities to distinguish the edible, dangerous, or fertile from the poisonous, safe, or infertile . . . Furthermore, we are able to distinguish real from apparent properties: our visual systems are good at tracking fixed and intrinsic color (i.e. surface reflectance) properties despite changes in appearance due to changes in lighting. (Bird 2007: 10)

Since our cognitive systems have evolved to distinguish real properties from merely apparent ones, it should not be surprising if the intuition that some properties are genuine while others are mere fabrications corresponds to a real distinction.

Second, sparse properties seem to be presupposed by a realist conception of science. When we do science, we are trying in part to discover rather than invent the properties that things have. Consequently, the practice of science lends itself to properties being sparse, since it is sparse properties that are liable to discovery as opposed to invention.

Bird's third argument is based on the idea that scientific terms get their referents fixed by their theoretical roles. The term 'mass' gets its referent fixed by its role in statements such as 'Force = mass × acceleration.' For any given term like 'mass,' it is possible to invent an indefinite number of terms that differ from it only slightly in

extension. The term 'mass*,' for instance, might include in its extension all the individuals to which 'mass' applies, minus this one or that one. Given that this is the case, we need to explain why it is that when scientists use the term 'mass' they manage to refer to mass instead of mass*. If properties are abundant, if they are just predicates or classes, then there is no easy way of explaining this. Since every predicate or class is ontologically on a par with any other, since there is nothing to set it apart ontologically from any other predicate or class, there seems to be no principled reason why 'mass' should refer to mass instead of mass*. But there is a principled reason if properties are sparse, for in that case mass is a property, whereas mass* is not. A framework committed to sparse properties is thus better suited for making sense of scientific practice than a framework committed to abundant properties.

Bird's final argument for sparse properties makes an analogous point about a scientific property's explanatory roles. The property of having mass but not the property of having mass* explains why a baseball accelerates at such-and-such a rate when it is batted with such-and-such a force, for it is mass and not mass* that factors into Newton's force law. But if mass and mass* are on a par with each other ontologically, as an abundant property view would have it, then there is no explanation for why mass plays this role and mass* does not. If properties are sparse, by contrast, then the explanation is clear: mass* is not a property, but mass is. Once again, therefore, sparse properties enable us to make sense of scientific practice in a way that abundant properties do not.

In addition to Bird's four arguments, there are two more. First, Peter Carruthers (2000: 35–7) advances an argument that appeals to change. Intuitively, we want to say that something undergoes a real change (as opposed to a mere Cambridge change (Geach 1969)) only when it gains or loses properties. Yet, this idea is difficult to accommodate if properties are abundant. Consider the abundant properties *being grue* and *being bleen*. An object is grue if it is observed to be green before, say, the stroke of New Year 2016, and otherwise is blue. And an object is bleen if it is observed to be blue before the stroke of New Year 2016, and otherwise is green. Suppose that we observe an emerald, *a*, to be green before New Year's Eve 2015. In that case, *a* is grue. Since it was not observed to be blue before the stroke of New Year 2016, it becomes bleen at that time. Yet, intuitively, we want to say that *a* underwent no real change. This intuition makes perfect sense if properties are sparse, since *a* underwent no change in its sparse properties. Since it does undergo a change in its abundant properties, the latter are ill-suited to accommodate our intuitions about real change. Consequently, if we expect a theory of properties to provide resources for explaining real change, we should take properties to be sparse, not abundant.

Finally, sparse properties are the price of admission to many philosophical debates, including debates in the philosophy of mind. The problem of mental causation, for instance, concerns whether mental properties can make any causal contribution to actions. The debate makes little sense if properties are taken to be predicates or

classes, for it is unclear in what sense either predicates or classes are candidates for making causal contributions to things. It seems, therefore, that anyone who wants to participate in these debates must suppose that properties are sparse.

The foregoing arguments give some reason for endorsing sparse properties instead of abundant ones. In addition, I'll argue in Chapter 11 that exponents of abundant properties have a difficult time formulating an acceptable definition of physicalism. Yet, despite everything I've said, many of the claims I intend to make can be formulated in a way that presupposes abundant properties. To my mind, the costs of doing so are undesirable; they include a good deal of metaphysical murkiness as well as terminological awkwardness, but dyed-in-the-wool proponents of abundant properties might think this is a price worth paying.

2.5 Conclusion

In this chapter, I've described and defended a basic substance-attribute ontology that is committed to individuals, properties, and states of affairs or events which are individuals having properties. It claims that individuals are the primary agents, and that they enter into causal relations on account of the properties they have. It is also committed to the Eleatic principle and sparse properties. The only individuals, it says, are ones that can enter into causal relations, and the only properties are ones that empower individuals to enter into those relations. This ontology is sufficient for some of our purposes in what follows, but for other purposes, including articulating an understanding of structure, we will need to have an ontology that is more developed. The goal of the next few chapters is to outline my *preferred ontology*. It develops the basic ontology in two ways. First, it takes properties to be tropes (Chapter 3). Second, it takes properties to be powers (Chapters 4 and 5).

3

Tropes

3.1 Tropes versus Universals

There are several ongoing debates about the metaphysics of properties (Armstrong 2005). An old one concerns whether properties are universals or particulars. Realists like Armstrong (1978a, b; 1983; 1989a; 1996a, b; 1997; 2005; 2010) claim that properties are universals. Universals are "repeatable" entities (Loux 2006): numerically one and the same universal can be instantiated by and thereby be wholly present in diverse individuals; it is a one-in-many, as Aristotle put it. Nominalists deny that there are universals. According to them, there are only particulars; although exactly what kinds of particulars exist is a matter of dispute. Extreme nominalists posit only individuals. Moderate nominalists posit other particulars besides. These particulars are often called 'abstract particulars' to distinguish them from individuals, which are the concrete particulars. Abstract particulars include classes (Quine 1960; Quinton 1957; Price 1953a, b; Lewis 1983b) and tropes (Stout 1921, 1923; Williams 1953, 1986; Campbell 1981, 1983, 1990; Simons 1994; Bacon 1995; Mertz 1996; Martin 1996a, b, 1997, 2007; Molnar 2003; Heil 2003, 2005; Mumford 2012).

Tropes are particularized properties, also called 'unit properties,' 'property instances,' 'individual accidents,' and 'modes,' among other things.[1] Intuitively, each is a *way* that an individual can be (Levinson 1978; Lowe 2006; Martin 2007). To understand the concept of a trope, it is helpful to contrast it with the concept of a universal. Consider two apples, *a* and *b*, which are qualitatively indistinguishable. The redness,

[1] There is an alternative way of understanding tropes which was introduced by Michael Loux (2015) and developed by Robert Garcia (2015; forthcoming). It claims not that tropes are particularized properties, but that they are propertied particulars. Loux describes the view as follows: "one might propose a nominalistic ontology that has as its metaphysical atoms what we might call 'tropers.' Whereas tropes are particular properties—things like this redness, this triangularity, this pallor, tropers are thin individuals—things like *this individual red thing*, *this individual triangular thing*, and *this individual pale thing*. The claim would be that familiar objects are bundles of compresent tropers. So the view would again dispense with properties and would insist that the ultimate constituents of familiar particulars are intrinsically characterized or natured, but would construe those constituents as particulars rather than universals. Such intrinsically characterized particulars would be the ultimate or underived sources of character: a familiar particular would be, say, pale because it has a pale troper as a constituent. On this view, the ordinary individuals of pedestrian acquaintance are bundles of tropes" (Loux 2015: 31). I look to defend a traditional trope theory, one that takes tropes to be particularized properties (what Garcia calls 'modifying tropes'), not tropers.

roundness, and other qualitative characteristics of *a* are completely indistinguishable from those of *b*. Since properties are supposed to account for the objective similarities and differences among things, these qualitative similarities should be explained by appeal to *a*'s and *b*'s properties. Realists explain them by claiming that there are universals. There is a universal, redness, for instance, that is instantiated by both *a* and *b*. Trope theorists deny this. There is not a single entity, they say, the universal redness, which is instantiated by both *a* and *b*; rather, *a*'s redness is a property that is numerically different from *b*'s redness. What explains the similarity between *a* and *b* if *a*'s redness and *b*'s redness are numerically different properties? The answer, say trope theorists, is that *a*'s redness and *b*'s redness exactly resemble each other. Because they exactly resemble each other, it can seem as though there must be an entity, a universal, which the two literally have in common. But according to trope theorists, this is not the case. Objective similarities can be explained by appeal to exactly resembling tropes. Saying that *a* and *b* have the same color is analogous to saying that a boy and his father have the same nose, or that two embarrassed celebrities arrived wearing the same dress. Statements like this do not posit a single nose or a single dress, but two exactly resembling ones. Realists insist that all similarities must be grounded in numerically identical universals, but trope theorists deny this. They take similarities as ground-level facts that stand in need of no further explanation.

Consider likewise how realists and trope theorists account for the reference of abstract nouns such as 'redness' or 'color' in statements such as 'Redness is a color.' Realists claim that these nouns refer to universals; trope theorists disagree. According to what is probably the most popular kind of trope theory, these nouns refer to classes of resembling tropes (Williams 1953, 1986; Campbell 1981, 1983, 1990; Martin 1996a, b, 1997, 2007; Molnar 2003; Heil 2003). When we say 'Redness is a color,' we are talking about the resemblance class of red tropes. There are different ways of constructing resemblance classes. One way bases them on handfuls of paradigmatic examples (Price 1953b: 20–3). In the case of redness, the exemplars might include stop signs and ripe tomatoes. The resemblance class based on these exemplars would include all and only those tropes that resemble each of the exemplars at least as closely as the exemplars resemble each other. On this view, saying that redness is a color amounts to saying that the resemblance class of red tropes is a subset of the resemblance class of color tropes.[2]

Many trope theorists have been attracted to bundle theories of substance (Stout 1921, 1923; Williams 1953, 1986; Campbell 1981, 1983, 1990; Simons 1994;

[2] An alternative worth exploring claims that the semantics of abstract nouns like 'redness' is actually similar to that of plural referring expressions like 'the tenors' in 'The tenors in the choir lost their pitch.' The term refers not to a class (the class that includes all and only the tenors in the choir); it refers instead to the individual tenors themselves. Trope theorists could claim that 'redness' has a similar semantics; it refers not to a class of tropes, but to the individual tropes themselves. Consequently, when we say 'redness is a color,' we mean that the various tropes in the extension of the term 'redness' are also tropes in the extension of the term 'color.' For more on plural reference, see Cameron (1999).

Mumford 2012), but I follow those trope theorists who remain committed to a substance-attribute ontology. They include Martin (1980, 1996a–b, 1997, 2007), Molnar (2003), Heil (2003, 2005), and trope theorists of the past such as Thomas Aquinas, William of Ockham, and perhaps even Aristotle.[3]

3.2 Arguments for Trope Theory

There are several arguments for trope theory. One kind of argument appeals to the particularity of causal relations. E. J. Lowe provides an example:

> [I]t seems that only particulars can participate in causal relationships and that an object participates in such relationships in different ways according to its different properties. Thus, it is a rock's *mass* that explains the depth of the depression it makes upon falling on to soft earth, whereas it is the rock's *shape* that explains the shape of the depression. Perception itself involves a causal relationship between the perceiver and the object perceived and we perceive an object by perceiving at least some of its properties—we perceive, for instance, a flower's colour and smell. But this seems to require that what we thus perceive are items that are unique to the object in question—*this* flower's redness and sweetness, say, as opposed to a universal redness and sweetness that are also exemplified by other, exactly resembling flowers. For, surely, in seeing and smelling this flower, I cannot be said to perceive the colour and smell of any other flower. (2006: 15)

Causal relations are relations between particulars. As Lowe indicates, this seems especially evident in the case of perception. My seeing this flower's redness involves a causal relationship between me and this flower. It is because of its redness—that property—that the flower is able to affect my visual organs the way it does. If it were not this flower's redness which affected my sensory organs, but something else, then there would be no straightforward account of what it means to say that I see this flower's redness. Trope theory thus emerges as the most straightforward way of accounting for causal relations.

A second kind of argument alleges that realism faces a variety of difficulties, and as a result, trope theory is relatively better off. One objection to realism, for instance, dates back to Peter Abelard, and even earlier to Boethius. It claims that realism is incoherent since it implies that one and the same entity, a universal, can have contradictory properties (Campbell 1990; Heil 2003; Molnar 2003). If, for instance, tomato_1 and tomato_2 both instantiate the universal redness, then that universal is both located where tomato_1 is and also not located there.[4] But surely, says the objection, nothing can have contradictory properties. Another objection concerns

[3] For an interpretation of Aristotle's metaphysics in terms of tropes, see Sellars (1957) and Irwin (1988). For an interpretation in terms of universals, see Loux (2008).

[4] This problem concerns a theory committed to immanent or Aristotelian universals, a theory that denies that there are uninstantiated universals. I accept Armstrong's (1978a, b; 1989a) reasons for favoring Aristotelian universals to Platonic ones.

the instantiation relation (Loux 2006: 31–2; Campbell 1990: 14–15; Price 1953b: 22–4; Donagan 1963: 135–9). The objection was first suggested by Plato in *Parmenides* (131E–132B ff.). Aristotle called it the problem of the 'Third Man' (*Metaphysics* 990b17). Its modern formulation is inspired principally by F. H. Bradley (1899). If realism is true, says the objection, then *a* is red because *a* instantiates the universal redness, for in general realism looks to explain statements of the form '*a* is *F*' by appeal to *a*'s instantiating the universal *F*. But consider now the statement '*a* instantiates the universal redness.' If realism is true, it seems that the truth of this statement must be explained the same way: it is true because *a* instantiates the instantiation of redness. But what explains the truth of this statement? Again, if realism is true, it seems that this must be true because *a* instantiates the instantiation of the instantiation of redness, and so on ad infinitum. Consequently, it seems that in order to explain the truth of a simple statement such as '*a* is red,' realists must posit an infinite number of instantiation relations. But surely this is absurd, says the objection. Among other things, it makes it difficult to see how the original statement '*a* is red' gets explained at all, for the first explanation in the series depends on the second, and the second depends on the third, and so on. Consequently, in order to grasp what makes it true that *a* is red, we would have to grasp every item in the infinite series. But we cannot do that. Hence, we cannot grasp the realist's explanation for what makes it true that *a* is red.

Objections of this sort are far from decisive. Realists have responses ready to hand. In response to the first objection, for instance, realists can reject the premise that one thing cannot be wholly present at two different locations (Lewis 1983b: 11). They can insist that this premise is based simply on intuitions that beg the question against realism by assuming that universals must obey the same kinds of ontological principles that apply to particulars. In addition, there are several ways that realists can respond to the second objection. They can posit instantiation as a primitive that does not give rise to a regress (Loux 2006: 35–6; Armstrong 1978a: 109ff.; 1997: 118–19).[5] They can deny that if the attempt to explain *a*'s being *F* generates further explananda (even an infinite number of them), this implies that *a*'s being *F* has not been adequately explained (Loux 2006: 33), or they can argue *tu quoque* that nominalist theories face analogous regresses (Russell 1912; Armstrong 1978a; Daly 1994; Macdonald 1998).[6] Nominalists have counter-responses, but the range of strategies and counter-strategies available both to them and to realists suggests that arguments like the foregoing result in stalemate.

[5] Armstrong has apparently not been satisfied with this approach, and describes a new alternative in Armstrong 2004b and 2005: 317–18. Some might regard it as a liability of his new tack that it brings his metaphysics closer to David Lewis'.

[6] Realists and trope theorists respond to this worry in similar ways. Both claim that the tie or nexus between an individual and one of its properties is not really a relation. Lowe (2006: 91–2), for instance, says that individuals are characterized by their tropes, and that characterization is not really a relation. As a result, he says, his account of tropes is able to avoid a Bradley regress.

Trope theorists can nevertheless still appeal to ontological parsimony to bolster their case (Campbell 1990; Bacon 1995). Nominalism, they can argue, is in general preferable to realism since it posits fewer basic types of entities: particulars only versus particulars plus universals. Ockham's razor bids us to posit no entities other than those needed to satisfy our theoretical demands, so all other things being equal, we should favor a coherent nominalism to a coherent realism. Moreover, trope theorists can say, among the available forms of nominalism, their own theory has arguably the most going for it (Armstrong 1989a: 119–31). We thus have good reason to accept trope theory. The crucial premise here is that all other things are in fact equal between realism and nominalism. What reason is there to think this is true? Armstrong (1989a: 122) suggests the following:

> [F]or each instantiated universal, a class of exactly corresponding tropes can be postulated as a substitute. The correspondence also goes the other way... So provided you abandon uninstantiated universals (good riddance, I say), and provided Universals theorists and Trope theorists coordinate their views on just what properties and relations the world contains, it is easy to pass back and forth between the theories. This is all rather nice business for the Trope theory... You get a construction that will do almost all the work that universals do, without having to postulate them. Paradise on the cheap! (Armstrong 1989a: 122)

If Armstrong is right, then resemblance classes of tropes can be taken to correspond to universals one-one. In that case, though, tropes and resemblance classes of them can be expected to do all the theoretical work of universals but with an ontology of only particulars. As a result, there seems good reason to favor trope theory on grounds of parsimony.

Martin (1996a: 72–3, 75–6) suggests a slightly different appeal to parsimony. I take it to be analogous to instrumentalist arguments against scientific realism. Instrumentalism and scientific realism both aim at empirical adequacy; both aim to account for the same body of empirical data, and to the extent that they succeed, they are both confirmed by that data. But, say instrumentalists, scientific realists go beyond what the data confirm when they insist that the predicates and terms of our best scientific theories must correspond to real entities. Martin suggests something analogous when it comes to tropes versus universals. Trope theory and realism about universals both aim at explaining the same facts about predication, similarity, and so on. To the extent that the theories succeed, those facts can be taken as confirming evidence in favor of them. Realists, however, go beyond merely accommodating the facts when they postulate universals, for the same facts can be accommodated by particulars alone. Consider again the similarity between *a*'s redness and *b*'s redness. According to trope theorists, *a* and *b* are similar to each other because they have exactly resembling tropes. This exact resemblance seems sufficient to accommodate the facts. Realists do not deny those facts; they do not deny that the redness of *a* and the redness of *b* are exactly similar. But realists appear to go beyond the facts when they claim that there is more than similarity but rather numerical sameness, the

sameness of numerically one and the same universal. This seems an unnecessary ontological extravagance. We thus have reason once again to favor a more frugal trope theory.

Molnar (2003) suggests a different argument for trope theory, one that resembles an inference to the best explanation. Trope theory, he argues, does the best job of blending the strengths of realism and the strengths of nominalism while avoiding their corresponding weaknesses:

I am convinced that there is something fundamentally correct in all versions of realism, and there is something (else) that is fundamentally correct in all versions of nominalism. It is desirable that trope theory should recover and preserve the insights of both realism and nominalism . . . *What is wrong in nominalism?* It seems perfectly reasonable to ask for a robust, ontologically grounded, explanation of the fact that a predicate applies to an object. Such explanations . . . typically present as explanans the existence of some properties borne by some objects . . . Nominalism, being globally anti-realist about properties, cannot offer any such explanations. Instead it restates the semantic criterion for the correct application of the predicate . . . This gives a formally adequate answer to the request for a truthmaker for the claim '*a* is *F*.' But it is not metaphysically adequate . . . [N]ominalists' well-founded distrust of universals misleads them into denying the reality of properties as such. *What is right in nominalism?* The great insight is particularism: everything is particular. Even the properties had by individuals . . . *What is wrong in realism?* Classic realism identifies properties with universals, which are strange posits indeed . . . I agree with the many philosophers who have thought that such entities cannot explain or cast light on anything. Whenever universals are invoked in an account of something . . . we understand less *after* the explanation is given than we understood before . . . *What is right in realism?* [R]ealism . . . is ontologically serious on an issue that calls for ontological seriousness. By including properties among the irreducible contents of this world, realism allows us to construct the robust explanations, of the facts of predication, of causation, of nomological connection, etc., that are blocked by nominalism. (Molnar 2003: 23–5)

Because tropes are particularized properties, trope theory combines the strength of realism with that of nominalism. It affirms the reality of properties while maintaining the particularity of everything that exists. It does this, moreover, without taking on either theory's liabilities. It provides a metaphysically robust framework for explaining predication, causation, and so on, yet it does so without positing universals.

A further advantage of trope theory is that it enables us to avoid positing states of affairs or events as a separate ontological category (Campbell 1981: 354–5; Williams 1986: 4; Armstrong 1989a: 117–19). This point becomes clear if we once again contrast trope theory with realism. Armstrong (1989a: 88–9; 1997: 115ff.) takes states of affairs to be a necessary addition to his ontology because individuals, universals, and the instantiation relation are insufficient by themselves to supply him with truthmakers for statements like 'The apple is red.' If apple *a* is red, there must be something in the world that makes it true that *a* is red. That something, it seems,

cannot be *a* by itself since it is possible for *a* to exist without being red, nor can it be the universal redness by itself since it is possible for that universal to exist without *a* instantiating it (perhaps it is instantiated by *b* instead). Nor can the truthmaker be the sum of *a*, redness, and the instantiation relation, Armstrong argues, since this does not amount to *a*'s being red. Consequently, says Armstrong, we should take *a*'s being red, the state of affairs itself, to be the truthmaker.

Trope theorists, however, do not need to posit states of affairs in this way, for on their view, individuals and properties (tropes) are together sufficient for truthmaking. What drives Armstrong to posit states of affairs is that on his view universals and the individuals that instantiate them are related only contingently: *a* need not instantiate *F*, and *F* need not be instantiated by *a*. According to at least one version of trope theory, however, the relation between a trope and its individual bearer is not contingent but necessary, something that trope theorists have termed 'nontransferability' (Heil 2003: 141–2; Molnar 2003: 43–6). Martin explains:

> Properties are nontransferable. The redness or sphericity of this tomato cannot migrate to another tomato. This is a consequence of the idea that properties are particular ways things are. The identity of a property—its being the property it is—is bound up with the identity of its possessor. (Martin 2007: 44)

Tropes belong necessarily to the individuals having them: *a*'s redness cannot belong to something other than *a*, any more than Eleanor's smile can belong to someone other than Eleanor. Consequently, there is no need to posit something in addition to tropes and individuals to tie one to the other and make it true that *a* is red. *a*'s being red—the trope itself—does the truthmaking work on its own.

In addition to the foregoing point, Bennett (1988: 90–1) argues that the best way of understanding a property exemplification theory of events is to construe events as tropes. Events with the identity conditions discussed in Section 2.1 must be either tropes or triples, Bennett argues, but it is implausible to suppose that events are triples, for many of the things we take to be true of events are not true of triples. For instance, events but not triples occur. Likewise, events are located in spacetime, but triples are not. Moreover, triples don't cause other triples, but events do cause other events. Consequently, Bennett concludes, events must be tropes. In addition, Campbell (1981: 354–5) argues that the theoretical jobs events are asked to perform in theories of causation are best performed by tropes. There are, then, good reasons to think that events are tropes. A corollary of this claim is that an ontology that includes events must be an ontology that includes tropes as well, or as Bennett puts it, "An enemy of tropes must either oppose events . . . contending there are no such things, or find a good rival account of what events are" (1988: 90).

A final reason to prefer a trope ontology is that it dovetails with an attractive view of powers, a point that I discuss in detail in Chapter 4. Before turning to that topic, however, I want to say a word about some objections to trope theory.

3.3 Objections to Trope Theory

The past three decades have made it increasingly clear that trope theory provides an account of properties that is at least as adequate as any other. Even Armstrong, who at one time dismissed trope theory (1978a), came to regard it as the most promising alternative to his own view:

> In my earlier work . . . I underestimated the strength of a tropes + resemblance (+ substance-attribute) view. In my present estimation . . . it is a close second to the first choice, which is a Realism about universals. (Armstrong 1989a: 120)[7]

Based on the foregoing discussion, moreover, it is not evident that tropes are in fact second best. We've seen that there are reasons to favor trope theory over Armstrong's preferred view. Are there any considerations that cut the other way?

Many criticisms of trope theory target the conjunction of tropes with a bundle theory of substance. Others target the idea that terms like 'red' and 'color' refer to resemblance classes of tropes, and yet others target the existence of tropes themselves. Because the ontology I've outlined rejects a bundle theory of substance, we can put objections of the first sort to one side, and focus on the other two.

Perhaps the most commonly heard objection to tropes claims that there cannot be brute resemblances among individuals. If *a*'s redness resembles *b*'s redness, then this resemblance must be explained by something. In particular, say realists, it must be explained by *a* and *b* instantiating the same universal. Trope theorists, however, reject the argument's premise; they take at least some resemblances among individuals to be brute. Realists might see the commitment to brute resemblances as a theoretical cost, but trope theorists take a commitment to universals to carry with it offsetting costs. Realists are likely to disagree that the costs are truly offsetting,[8] but the replies and counter-replies available to trope theorists and realists alike ought to make us suspect that debate over this objection will end in stalemate. It is perhaps for this reason that ardent critics of trope theory, such as David Armstrong, have advanced different arguments.

Armstrong (2004a: 43–4; 2005: 310) has objected to tropes by appeal to an argument originally suggested by Herbert Hochberg (2001a: 69–70; 2001b: 178–9). The gist of the objection is that trope theorists ask tropes to play too many theoretical roles. Consider *a*'s redness, a trope that is ostensibly simple, not complex. Given this trope's simplicity, we would expect that it would not be able to play very many theoretical roles, yet this is exactly what trope theorists demand of it. According to

[7] Armstrong's realism is, in fact, very close to trope theory since it rejects uninstantiated universals and claims that universals are wholly exhausted by their instances (Martin 1993: 178–9). Armstrong (1989a: 122; 1996b: 88) and Martin (1996a: 72–3, 75–6) have even suggested that their views (the former committed to immanent universals, the latter to tropes) might be merely notational variants of each other.

[8] Pickel and Mantegani (2012) argue, for instance, that nominalist theories that are reticent about explaining why some things resemble each other carry a heavier ontological burden than realism.

them, *a*'s redness must have a nature that distinguishes it from every other trope, that makes it exactly similar to some of them, that makes it less similar to others, and that makes it quite dissimilar to yet others. But it is implausible to think that a simple trope could perform all these theoretical roles, says Armstrong: "You would expect that to the complexity of the various truths made true by the tropes, there would need to be some complexity in the tropes themselves" (2005: 310).

The Armstrong-Hochberg objection is based on the premise that tropes can be simple. But why should we accept this premise? Armstrong defends it by appeal to the simplicity of properties: "Presumably properties can be simple, so there can be simple tropes" (2005: 310). But this argument appears to be a non sequitur that equivocates on 'simple.' Properties might be considered simple in a number of respects. Monadic properties might be considered simple by virtue of having only one term; qualitative properties might be considered simple by virtue of there being no analysis of what it takes for something to have them. Properties can be considered simple in these and many other respects, yet these respects are still compatible with each property playing a variety of theoretical roles—with each grounding similarity, grounding numerical difference, being a causal enabler, being a causal explainer, and so on. Armstrong disagrees. This is not surprising since Armstrong endorses a view of properties that divorces them from most of their theoretical roles. As we'll see in detail in Chapter 5, according to Armstrong, the natures of properties consist simply in primitive principles of identity, what some philosophers have called 'quiddities' (Black 2000). Any other theoretical roles properties might be expected to play are outsourced to other things such as laws of nature. It should not surprise us, then, that Armstrong balks at the suggestion that properties should play a variety of theoretical roles. What remains unclear is why we should favor his understanding of properties to an alternative that takes a variety of theoretical roles to be integral to what they are. In fact, we'll see in Chapter 5 that Armstrong's view of properties has some rather unpalatable implications.

A third objection to trope theory appeals to a version of Bradley's regress analogous to the one facing realists. Recall that according to realists, a particular *a* is red because *a* instantiates the universal redness. A Bradley regress looms because it seems that realists are committed to saying that *a* instantiates the universal redness because *a* instantiates the instantiation of redness, and this is true because *a* instantiates the instantiation of the instantiation of redness, and so on ad infinitum. According to critics, trope theory faces an analogous problem. The relation between a trope and its bearer is not instantiation, which is a relation between a universal and a particular; it is something else. Lowe (2006) calls it 'characterization.' Individuals are characterized by their modes (Lowe's preferred term for tropes). But, says the objection, whatever trope theorists decide to call the relation between a trope and its bearer, they face the same problem as realists: *a* is characterized by *a*'s redness because *a*'s redness is characterized by the characterization of *a*'s redness, which is characterized by the characterization of the characterization of *a*'s redness, and so on ad infinitum.

In response to this objection, trope theorists can respond in a manner analogous to those realists who claim that instantiation is a sui generis relation that does not need to be explained by appeal to further instantiation relations. Lowe articulates the response in the following way:

The flower's particular redness is a mode—a particular way that flower is—and one which for that reason may be said to 'characterize' the flower. Quite literally, the mode is a particular 'characteristic' of the flower . . . I am reluctant to say that characterization is a *relation* between a particular thing and its modes. For then, it seems, we should have to conceive of a thing and one of its modes as being the relata of a further *relational* mode, which would in turn 'characterize' . . . those two relata . . . And it is easy to see that in this way an infinite regress would be generated . . . [W]e can draw comfort from [the] observation that not every meaningful predicate need be supposed to denote a property—or, in this case, a relation. Just because 'is characterized by' is a meaningful relational predicate . . . we need not conclude that that predicate denotes a relation in which the flower and its particular redness stand to one another. (2006: 92)

If characterization, as Lowe calls it, is not a relation, then trope theorists avoid a Bradley regress. Moreover, because tropes are nontransferable—because *a*'s redness cannot belong to anything other than *a*—it is prima facie more plausible for a trope theorist to respond in this way than a realist. Because the universal redness can be instantiated by things other than *a*, it is easier to view instantiation as a relation between distinct entities. As a result, realists have a heavier burden to carry when making this kind of response.

A fourth objection to tropes has been advanced by Cynthia Macdonald (1998: 346–7). It concerns the implications of trope theory for the philosophy of mind. Macdonald argues that realism is superior to trope theory because trope theory renders the notion of nonreductive monism incoherent:

Non-reductive monism is the view that each mental event is a physical event although mental properties are neither reducible to nor correlated in a . . . lawlike way to physical ones. [T]his theory seems . . . to reconcile monism at the level of particular events and their causal transactions, with the *sui generis* distinctness of the mental and physical at the level of properties. However, trope theory has difficulties providing the underlying metaphysics for such a view . . . How . . . are we to understand the claim that each mental event is a physical event? Suppose . . . that we take it to mean that this pain-trope just is this neurophysiological event-trope. Then the distinction between mental and physical properties seems unsustainable . . . For properties are classes of exactly resembling tropes, and physical tropes that are exactly resembling will thereby be mental tropes that are exactly resembling . . . [T]here will be no means by which to distinguish mental properties from physical properties. (Macdonald 1998: 346)

Since realism does not have the same result, Macdonald argues, realism is superior to trope theory.

Macdonald's argument tacitly assumes that nonreductive monism is committed to property dualism, the claim that mental and physical properties are distinct ("distinctness of the mental and physical at the level of properties," as she puts it).

I will argue in Section 11.4 that a characterization of nonreductive monism along these lines cannot be sustained. Roughly, the argument says that monism of any sort implies that everything is of only one kind. Physical monism, for instance, claims that everything is physical. Given the basic ontology of individuals, properties, and events sketched earlier, this implies that all individuals, properties, and events must be physical. With these terms in place, a physical monist cannot coherently claim that mental and physical properties are distinct, for if mental properties exist at all, this claim implies that they must be nonphysical. But if there are nonphysical properties, then not everything is physical, contrary to physical monism. The correct way of characterizing nonreductive monism is not in terms of a dualism of properties, but in terms of a dualism of descriptive and explanatory resources or interests. Davidson expresses the idea in terms of our commitments to the different internal standards of psychological and physical discourse, respectively:

> When we turn to the task of interpreting the pattern [of verbal behavior] we notice the need to find it in accord . . . with standards of rationality . . . [T]he case is no different with beliefs, desires, and actions . . . As long as it is behavior and not something else we want to explain and describe, we must warp the evidence to fit this frame. Physical concepts have different constitutive elements. Standing ready, as we must, to adjust psychological terms to one set of standards and physical terms to another, we know that we cannot insist on a sharp and law-like connection between them . . . The limit thus placed on the social sciences is not set by nature, but by us when we decide to view men as rational agents with goals and purposes, and as subject to moral evaluation. (Davidson 1974: 239)

According to Davidson, psychological discourse is irreducible to physical theory, not because there are two different kinds of properties, physical and mental, but because there are two different kinds of conceptual schemes with "disparate commitments" (Davidson 1970: 222). On this view, the 'nonreductive' in 'nonreductive monism' derives not from a nonmonistic metaphysics that endorses property dualism, but from the plurality of our descriptive and explanatory resources.

If the gist of this argument is correct, then Macdonald's argument against trope theory fails. She mentions several other arguments against tropes, but those arguments concern the conjunction of trope theory with bundles. She does not consider the conjunction of tropes with a substance-attribute ontology of the sort defended here. Admittedly, a substance-attribute ontology loses some of its edge when it comes to Ockham's razor since it posits individuals and properties instead of just properties, but it is also spared the problems facing tropes plus bundles (Martin 1980: 7–8; Armstrong 1989a: 115; Molnar 2003: 47–54).

In addition to the foregoing objections, David Manley (2002) advances two arguments to the effect that properties cannot be resemblance classes of tropes. The first argues that a trope theory committed to resemblance classes faces a version of the coextension problem discussed in Section 2.3. Imagine a possible world populated by only red objects. If properties are resemblance classes of tropes, then

redness is the same property as coloredness in the red world since the class of red objects is the same as the class of colored objects. Presumably, the names of properties such as ‘redness’ and ‘coloredness’ are rigid designators; each designates the same property in every possible world in which that property exists. Consequently, if redness is coloredness in the red world, it follows that redness must be coloredness in all possible worlds. But redness is not coloredness in all possible worlds since there are actually colored objects that aren’t red. Consequently, the objection concludes, properties must not be resemblance classes of tropes.

The thing to say in response to this objection, I think, is that the trope theory I’ve articulated is not committed to claiming that resemblance classes of tropes are properties. What it is committed to claiming is rather that properties are tropes themselves. Classes of tropes, by contrast, are just that: classes. For reasons discussed in Chapter 2, if properties are sparse, they cannot be classes since classes don’t confer powers on individuals as sparse properties do. Consequently, resemblance classes of tropes are not even candidates for being properties on the view of tropes I’ve articulated.

Realists might think that this response gives up the game, that admitting that terms like ‘redness’ do not refer to properties is tantamount to admitting defeat. This attitude assumes that a workable theory of properties must include determinable properties such as redness and coloredness. Yet, there is reason to think that there are no determinable properties, but only determinable predicates.

To appreciate this idea, recall first that according to the account of properties defended in Chapter 2, properties are sparse. This account implies that any given predicate need not express a genuine property. In order not to beg any questions, then, let us speak not of determinable properties, but rather of determinable predicates such as ‘is red’ or ‘is colored.’ The claim that there are no determinable properties follows from two premises: first, determinable predicates are higher-order predicates; and second, higher-order predicates do not express genuine properties. I’ll discuss these premises in order.

Higher-order predicates are logical constructions with definitions that quantify over properties. Suppose, for instance, that $F_1, F_2, \ldots, F_n$ are properties, and that we define being H (or the predicate ‘is H’) as follows: necessarily, for any x, x has H if and only if x has *some* F-property or other, either F_1, or F_2, or . . . , or F_n. Because H’s definition quantifies over other properties, ‘is H’ is a higher-order predicate (and philosophers who endorse abundant properties would say that being H is a higher-order property). The general logic of determinable predicates seems to imply that they are higher-order predicates. Suppose that ‘is D^*’ is a determinable predicate, and that ‘is D_1,’ ‘is D_2,’ . . . , ‘is D_n’ are its corresponding determinate predicates. In that case, the following seem to be true of D^* and the D_is:

(1) Necessarily, for any x, if x is D^*, then there is a D_i such that x is D_i.
(2) Necessarily, for any x, if there is a D_i such that x is D_i, then x is D^*.

(3) Necessarily, for any x, if there is a D_i such that x is D_i, then possibly, there is a y and a D_j such that y is D_j, where $j \neq i$.
(4) Necessarily, for any x, if there is a D_i such that x is D_i, then x is not D_j, for any $j \neq i$.[9]

The conditions that interest us here are (1) and (2). Consider a determinable predicate such as 'is red' and the predicates 'is R_1,' 'is R_2,' . . . , 'is R_n' which express all the determinate shades of red there are: Mandan red, cherry red, and so on. According to (1), necessarily, if something is red, then it is some determinate shade of red or other, either R_1, or R_2, or . . . , or R_n. According to (2), on the other hand, necessarily, if something is a determinate shade of red, then it is red. The necessity in both cases appears to be conceptual; it is a conceptual fact that if a determinable predicate applies to something, then one of the corresponding determinate predicates applies to that thing as well, and conversely, if a determinate predicate applies to something, then so does the corresponding determinable predicate. In that case, however, it seems that determinable predicates are higher-order predicates. Necessarily, something satisfies a determinable predicate if and only if it satisfies some corresponding determinate predicate. Something is colored if and only if it is some more determinate shade of color; something is a mammal if and only if it is a member of some more determinate animal kind; something is tall if and only if it is some more determinate height, and so on. We have some reason to think, then, that determinable predicates are all higher-order predicates.

In Section 5.3, I discuss in detail an argument to the effect that there are no higher-order properties, but only higher-order predicates. If they express any properties at all, higher-order predicates express the properties over which their definitions quantify. For instance, suppose that something has H if and only if it has some F-property or other, either F_1, or F_2, or . . . , or F_n. Suppose, moreover, that an object a satisfies the predicate 'is H' on account of having the property F_1. In that case, the predicate 'is H' expresses the property F_1. What it does not express in this case or any other is a higher-order property, H, that is distinct from the various F_is. There are no higher-order properties, then, no properties with definitions that quantify over other properties; there are only higher-order predicates or concepts or descriptions (Kim 1998: 104; Heil 2003: 45). If this is the case, then determinable predicates do not correspond to any properties distinct from the determinate properties that satisfy them. There are no determinable properties. If the object a satisfies the predicate 'is red' on account of being Mandan red, then in this particular case, the predicate 'is red' expresses the fully determinate property of Mandan redness. If b satisfies the predicate 'is red' on account of being cherry red, then in this particular case, the

[9] These correspond to the conditions on determinable predicates in Jaworski 2009: 142, which are gleaned from Prior's (1949), Johnson's (1964), and Yablo's (1992) discussions.

predicate 'is red' expresses the fully determinate property of cherry redness, and the same is true of all determinable predicates.

The upshot of these considerations is that there are no properties such as redness or coloredness in general; there are only fully determinate properties, which are tropes. If terms like 'redness' or 'coloredness' refer to anything at all, they refer to resemblance classes of fully determinate tropes, and there is nothing odd about supposing that in some possible worlds the class of fully determinate color tropes may be coextensive with the class of fully determinate red tropes.

Manley's (2002) second argument claims that a trope theory committed to resemblance classes faces a version of what Goodman (1966) calls the problem of 'imperfect community.' Let *w* be a possible world with only three objects: *a*, *b*, and *c*. *a* is pink, *b* is baby blue, and *c* is magenta. Suppose, moreover, that pink and magenta are both reddish, that magenta and baby blue are both blueish, and that pink and baby blue are both pale. Intuitively, we want to say that the class of reddish tropes includes only *a*'s being pink and *c*'s being magenta. The problem is that *a*'s being pink resembles *b*'s being baby blue to the same degree that it resembles *c*'s being magenta. As a result, it appears that *b*'s being baby blue must be included in the same resemblance class as *a*'s being pink and *c*'s being magenta. But intuitively, a resemblance class that comprises all three tropes does not correspond to any natural property. Hence, resemblance classes of tropes cannot be natural properties.

Here, I think, the response to the objection has two parts. The first we have considered already: the trope theory I've articulated is not committed to resemblance classes being properties; in fact, as we've seen, it rejects that claim. Second, resemblance classes come in degrees of naturalness just as classes of particulars do (Lewis 1983b; 1994). We are capable of coining predicates ad libitum, including predicates that apply to tropes that only loosely resemble each other, or that resemble each other in ways that do not satisfy intuitive notions of naturalness. But the existence of nonnatural classes of tropes should not count as a strike against a trope theory unless that theory is committed to resemblance classes being properties with high degrees of naturalness. Since the trope theory I've articulated denies that resemblance classes of tropes are properties, it should have nothing to fear from the objection.

Critics might retort that the responses I've advanced to the foregoing objections (eschewing determinable properties and denying that resemblance classes of tropes are properties) exact a price that isn't worth paying. To philosophers of this mindset, I would offer a compromise: a theory like Lowe's (2006) that posits universals in addition to tropes. On this view, redness and coloredness are universals instantiated by tropes. These universals are what terms like 'redness' and 'coloredness' refer to, and they are what explain the high degree of naturalness enjoyed by some classes of tropes and not others. Nothing I want to say about hylomorphic structure later on depends on denying that there are universals that play these theoretical roles. So philosophers who are uncomfortable with the responses I've made to the coextension problem and the problem of imperfect community can make use of universals.

3.4 Conclusion

I've argued in favor of trope theory over realism, and have defended it from some objections. Many of the arguments that would seem to favor tropes over universals or universals over tropes result in stalemate, but there is still some reason to prefer tropes on grounds of ontological parsimony and theoretical advantage. In addition, trope theory dovetails with an attractive view of powers, something I discuss in detail in Chapter 4. Having said all this, the account of hylomorphic structure I develop in later chapters is compatible with a view like E. J. Lowe's (2006), which posits universals in addition to tropes to help account for naturalness and abstract reference. My own preferred ontology does not posit universals, but I needn't insist on rejecting them in order to accomplish my theoretical aims.

4

The Identity Theory of Powers

4.1 One Property, Many Roles

This chapter and the next continue the task of developing a workable metaphysics for understanding hylomorphic structure. I don't purport that it is the only workable metaphysics suited to this purpose. I don't even purport that it is the best. There may be other metaphysical frameworks that would provide a better foundation for the hylomorphic theory I'm looking to construct. My contention is simply that the metaphysics I've been articulating is defensible, and that it provides a workable basis for a hylomorphic account of structure. That metaphysics is committed to properties being powers.

The basic ontology described in Chapter 2 says that properties are entities that confer powers on individuals. There is a debate in metaphysics about what exactly this claim means. The view I favor is a version of one defended by C. B. Martin (1993, 1996a, b, 1997, 2007), John Heil (2003, 2005), and Martin and Heil (1998, 1999). Heil calls it the 'identity theory.' I will adopt his terminology, but to avoid any confusion with the psychophysical identity theory, I will call it the *identity theory of powers*.[1]

The identity theory of powers claims that one and the same property plays a variety of theoretical roles which we express using different vocabularies. Sometimes we use a dispositional vocabulary, other times we use a nondispositional one. These different vocabularies create the impression that there are different kinds of properties: dispositional and categorical (or qualitative). According to the identity theory, though, these vocabularies describe the very same properties; they just bring out the different theoretical roles these properties play.

The identity theory of powers claims that each property is essentially dispositional, and in this sense it is similar to pure dispositionalist theories (Shoemaker 1980; Mumford 2004; Bird 2007). Each property essentially empowers its individual possessor to interact with other individuals in various kinds of ways. A diamond's hardness empowers the diamond to do a variety of things—to scratch glass, for instance. It is essential to the hardness that it empowers the diamond to do these things; it plays this power-conferring role in every possible world in which it exists.

[1] There are different interpretations of the identity theory of powers. The one I defend differs significantly from Engelhard's (2010), and is closer to the interpretation defended by Jonathan Jacobs (2011).

We describe this role in a variety of ways. We say that the diamond is able (or has the power or potential, or capacity) to scratch glass, or that the diamond would scratch that mirror if raked across its surface. But the diamond's hardness plays other roles that we describe in different terms. We say, for instance, that the diamond has a tetrahedral arrangement of carbon atoms. According to the identity theory of powers, these descriptions are of numerically one and the same property. The diamond's hardness = the diamond's power to scratch glass = the diamond's having a tetrahedral arrangement of carbon atoms. These descriptions simply bring out different theoretical roles that the one property plays. On the one hand, properties are tropes or modes: they are *ways* that individuals are (Lowe 2006; Engelhard 2010). Nondispositional descriptions such as 'The diamond is hard' express these ways. On the other hand, because properties are sparse, they are causal enablers: they empower individuals to enter into causal relations with each other. Dispositional descriptions such as 'The diamond would scratch that mirror if raked across its surface' express how an individual's properties empower it.[2] According to the identity theory, moreover, every power (with a possible exception I'll discuss momentarily) is both a power and a stable manifestation of a further power or powers. Nondispositional descriptions such as 'The diamond has a tetrahedral arrangement of carbon atoms' express a property's status as a stable manifestation—an actualization, we might call it—of another power or powers, in this case, the power the carbon atoms have to be arranged tetrahedrally. One property is thus simultaneously a stable manifestation of a power (or powers) and a power itself, both an actuality and a potentiality.

Martin and Heil mention the identity theory's Lockean pedigree, but to my mind the more obvious historical antecedent is Aristotle. Almost everything on Aristotle's view is both an actuality of some potentiality and a potentiality for some further actuality. The only exceptions are prime matter and God. On Aristotle's view, the elements earth, air, fire, and water are the most basic identifiable kinds of stuff. Aristotle nevertheless maintains that the elements can be transformed into one another. Given his view of change, this implies that each must be the actuality of some further potentiality; there must be something which has the power to become earth or air, fire or water. Since earth, air, fire, and water are the most basic identifiable kinds of stuff, however, that something cannot be a further identifiable kind of stuff, nor can it be characterized by the possession of any properties since properties, for Aristotle, must be possessed by substances, and substances must

[2] Engelhard (2010) grounds the dispositional/nondispositional distinction in an ontological distinction between universals and tropes since she endorses an ontology like Lowe's (2006). I prefer an account that takes the distinction to be purely conceptual. *a*'s being *F* can be conceived as a way the individual *a* is, but it can also be conceived as enabling *a* to stand in certain causal relations. Engelhard follows Lowe in taking dispositional descriptions to correspond to universals. A dispositional description expresses how, say, *a*'s being *F* enables *a* to stand in causal relations that are exactly similar to individuals with exactly similar properties. Unlike Engelhard and Lowe, however, I don't think an identity theory of powers need be committed to saying that every dispositional description has a universal cast.

belong to kinds. As a result, the something which has the power to become earth or air, fire or water—what Aristotle calls 'prime matter'—can only be characterized in terms of its role as a potentiality for becoming one of the basic identifiable kinds of stuff. It is a pure potentiality which is posited by drawing an analogy with other transformations that conform to the familiar actuality-potentiality pattern of explanation. An analogous point is true of God on Aristotle's view. There are no unactualized potentialities in God. God is pure actuality. As a result, God cannot be characterized in terms of being or having the potentiality for some further actuality. Prime matter and God are thus the limit cases of pure potentiality and pure actuality on a continuum in which otherwise every actuality is a potentiality and every potentiality an actuality. Interestingly, Martin calls his own view the 'Limit View' for analogous reasons. If there is a difference between him and Aristotle on this point, it is that Aristotle takes the limit cases to be real, whereas Martin takes them to be merely abstract postulates:

> To speak of a qualitative property is to take some real property as *only* at its bare potency-free purely qualitative limit, which of course, it never is. To speak of a dispositional property is to take some real property as *only* at its purely dispositional non-qualitative limit which, of course, it never is. No real property of an object, event, process or even space-time segment or field can be thought of as existing at *either* limit. The thought of anything being at either the limit of the purely and only qualitative disposition-free pure act of being . . . or the limit of the pure state of potency . . . is conceptual artifice and unrealizable abstraction. (1996a: 74–5; cf. Martin 1997: 215)

For our purposes, we needn't take a stand on whether or not the limits of pure potentiality and pure actuality can actually exist. The nonlimit cases are the ones that interest us, and on these cases, Aristotle and Martin agree that numerically one and the same property plays different theoretical roles which we come to grasp through abstraction.

Many of the claims I discuss in this chapter and the next have been defended in one way or another by Martin and Heil, but it is worth noting at the outset two differences between the way they develop the identity theory of powers and the way I do. The current discussion of powers in the literature has been framed as a debate about whether properties are fundamentally categorical or fundamentally dispositional. Properties such as roundness and redness are taken to be paradigmatically categorical. Properties such as being able to roll down an inclined plane or being able to reflect light with a 750nm wavelength, by contrast, are taken to be dispositional since they support counterfactuals such as 'If *a* were placed on an inclined plane, *a* would roll down it' and 'If *a* were exposed to light with a 750nm wavelength, then *a* would reflect that light.' Martin (1996a) eschews talk of categorical properties on the grounds that calling properties 'categorical' covertly begs the question against an identity theory of powers. The identity theory claims that so-called categorical properties are identical to dispositional properties, yet the term

'categorical' is so imbued with nondispositional connotations that its use suggests wrongly that an identity theory is ruled out a priori. Martin's preferred term is 'qualitative,' and Heil (2003, 2005) follows his lead. In describing the identity theory, I have avoided speaking either of categorical properties or of qualitative properties. I agree with Martin that the term 'categorical' is prejudicial. A commitment to categorical properties implies a specific metaphysics of properties that rules out alternatives like the identity theory of powers—a point I discuss in greater detail in Chapter 5. But I have also avoided characterizing properties as qualitative. The reason is that I think the term 'qualitative' is overcommitted as well.

The notion of a property has a very general scope. The notion of a quality has a much narrower one since it is linked to our experiences of things. It is because of this experience-linkage that Heil argues that rejecting the identification of powers with qualities veers toward idealism: "If minds have qualities but no material thing has qualities, then minds are not material things" (2005: 351). It is not evident to me that defenders of the identity theory of powers need to commit themselves to the claim that every property is characterizable in experience-linked terms. In fact, Molnar argues that not all properties can be thus characterized:

> [A]t the quantum level of nature, there are no . . . qualitative properties . . . We accept the existence of qualities because acceptance is *forced upon us* in experience. One needs but to recall . . . the *intrusiveness* of what is qualitative in perception and sensation to see that there is nothing 'theoretical' about the qualities disclosed in experience . . . By contrast, physical qualia are not items in good ontological standing . . . First, at the level of the fundamental constituents of matter, we are dealing exclusively with sub-observables. No qualities of the fundamental particles are given to us in experience . . . Second, any qualities we might postulate for the particles . . . are explanatorily idle. The only intrinsic properties needed to explain the behavior of the electron are its powers . . . If the electron had any qualities, they would not add value to the physical explanations we can now obtain by reference solely to its dispositions. (2003: 178)

If Molnar is right, then the notion of a property outstrips the notion of a quality. As a result, characterizing properties as qualitative, as Martin and Heil do, suggests an unwarranted empirical commitment to the claim that all properties must be given to us somehow in experience. I think identity theorists do well to avoid this commitment. I've suggested that they say instead that there are different vocabularies, some dispositional, others nondispositional, for describing the different theoretical roles that properties play. Among the roles that we use nondispositional descriptions to express are the statuses of properties as stable manifestations or actualizations. Since the notion of an actualization has a broader scope than the notion of a quality, identity theorists can avoid the liability Molnar's argument brings out in Martin and Heil's way of articulating the view.

A second difference between Martin and Heil's way of articulating the identity theory of powers and mine is that they reject talk of levels in nature, whereas I do not. Later, in Section 6.6, I will argue in detail that their grounds for rejecting levels are

misguided. Heil's (2003: chapter 5) argument in particular conflates the notion of higher levels with the notion of higher orders. The argument might succeed in showing that there are no higher-order properties (a claim with which I agree—see Section 5.3), but there are other ways of understanding levels in nature that are not committed to there being higher orders. In Chapter 6, I describe one such understanding based on the notion of composition.

4.2 Directedness and Mutual Manifestation

The identity theory of powers has several noteworthy features. First, it claims that powers are essentially directed toward their manifestations. Fragility is essentially directed toward breaking, solubility toward dissolving, and so on. This directedness is a primitive feature of powers. Under the right conditions, empowered individuals manifest their powers by acting in the ways toward which their powers are directed: the fragile objects break, soluble materials dissolve. And that's that: nothing further is required for empowered objects to act. They are simply "ready to go," as Martin (2007) says.

Some philosophers have drawn analogies between the directedness of powers and the directedness of intentional mental states (Martin and Pfeifer 1986; Place 1996a, b; Molnar 2003).[3] Intentional mental states are said to be directed at things. My desire is essentially a desire *for* something, my fear is essentially a fear *of* something, and so on. Something analogous is true of powers: they are essentially powers *for* various exercises or manifestations. Likewise, just as my desire can remain unfulfilled and my fear unrealized, so too a power can remain unmanifested. A quantity of table salt has the power to dissolve in water; that power actually exists, yet it might never actually be manifested. It is possible that the salt might remain forever undissolved. Martin (1996a) defends this idea with an example: it seems possible that there might be fundamental physical particles in the universe that have the power to interact in various ways with particles around here, and yet that are so far away that they reside outside the light cone of the particles around here. The two groups of particles thus never actually interact, yet it seems obvious that the distant particles still have the power to interact with the local ones.

If the directedness of powers is analogous to the directedness of intentional mental states in the ways just described, then identity theorists can avoid a charge that has

[3] Place takes the directedness of intentional states to be sufficient for something to have intentionality, and he takes intentionality to be the mark not of the mental specifically, as Brentano thought, but of the dispositional. Mumford (1999) resists this conclusion since he thinks it leads to animist or panpsychist views, a worry expressed by Martin and Pfeifer (1986). Bird (2007: 114–26) argues against Place that intentionality has other characteristics which Place either wrongly dismisses or else doesn't consider at all. It is true, for instance, that dispositions might be directed at their manifestations in a way that resembles the directedness of intentional states, and also true that the manifestations at which dispositions are directed need not actually exist, but intentional states are also extrinsic and indeterminate in a way dispositions are not. As a result, Place's conclusion does not follow.

sometimes been levied against theories of powers, namely that they imply a commitment to Meinongian nonactual entities (Armstrong 2002). The Meinongian worry arises on account of two premises: (1) powers are relations to their manifestations, and (2) powers need not be actually manifested. Relations are extensional. In order for *a* to stand in a relation to *b*, *a* and *b* must both exist. Consequently, if powers are understood as relations to their manifestations, as Premise (1) would have it, then powers and their manifestations must both exist. But according to Premise (2), a power need not be actually manifested; it is possible for the table salt to remain forever undissolved or for the distant and nearby particles never to interact. Since powers must be related to their manifestations in order to exist, it follows that there must be nonactual entities to stand in relation to unmanifested powers. Hence, the argument concludes, the identity theory is committed to Meinongian nonactual entities. In response to this argument, identity theorists reject Premise (1). Powers are not real relations, they say; they are instead to be understood in the way described above, a way analogous to intentional mental states; they have a directedness that does not depend on the existence of what they are directed toward.

It's important to note that the analogy with intentional mental states is merely that: an analogy. It would be a mistake to think that the directedness of powers is literally a species of intentionality (Bird 2007: 118–26). If anything, it is the other way round: intentional mental states are powers and the directedness of those states is a species of the directedness of powers in general. But an identity theorist needn't be committed to this claim either.

Another feature of the identity theory is that powers are manifested only in specific circumstances and only in conjunction with individuals that have reciprocal powers—what Martin calls 'reciprocal disposition partners.'[4] Powers can be manifested both actively and passively, both in the ways individuals affect things and in ways they are affected by them.[5] Powers are manifested if and only if individuals with reciprocal powers are conjoined in the right circumstances. Water, for instance, manifests its power to dissolve things only in conjunction with things that have the power to be dissolved by it. Moreover, when water is conjoined in the right circumstances with something that is water soluble, both simultaneously manifest their reciprocal powers: the water dissolves a quantity of salt and a quantity of salt is dissolved by the water.[6]

Harré and Madden's (1975) examples of radioactive decay and ammonium triiodide seem initially to provide counterexamples to the general rule that powers are manifested or exercised only in pairs, or triples, or . . . *n*-tuples. But even here it might be possible to understand the cases in a way that conforms to the general reciprocity model. At the very least, the environment surrounding the radioactive nuclei or the

[4] Other defenders of this thesis include Mumford and Anjum (2011), and Marmodoro (forthcoming).

[5] This should not be confused with the claim that powers come in both active and passive varieties. See Marmodoro (forthcoming) for more on this point.

[6] Marmodoro (forthcoming) has a similar view of the reciprocal manifestation of powers.

ammonium tri-iodide cannot include any agents that inhibit the exercise of their powers to decay or explode, respectively. Environments that are free of inhibitory factors might then be viewed as reciprocal disposition partners for the decaying nuclei and the explosive compound.

In addition, the exercise of some powers can inhibit or excite, impair or enhance, strengthen or weaken the exercise of others. An antidote has the power to inhibit the power of a poison, and there might be other things which have the power to enhance it. These observations reflect a more general point: the same power can manifest itself differently in conjunction with different disposition partners. To use Heil's example: a ball will roll on a hard surface on account of its roundness, and it will make a concave depression in a soft surface on account of that same roundness. The same property, the ball's roundness, manifests itself in different ways in conjunction with different disposition partners.

4.3 Identity Conditions for Powers

The foregoing remarks provide a preliminary basis for formulating identity conditions for powers. The idea that powers can be identified by their manifestations seems plausible prima facie, but as the remarks in Section 4.2 indicate, the identity theory of powers is committed to rejecting the thesis that every power corresponds to exactly one manifestation. It claims instead that a single power is capable of manifesting itself in different ways in conjunction with different disposition partners. It nevertheless seems possible to elaborate the basic idea in terms of the principle that two powers cannot manifest themselves in exactly the same ways in all possible circumstances. If Power_1 and Power_2 are truly different, there is some possible circumstance in which Power_1 and Power_2 will manifest themselves differently.

Suppose that *a*'s having P_1 and *b*'s having P_2 are reciprocal powers that manifest themselves in way M_1 when conjoined under the right conditions, C_i. Every power comprises a range of possible manifestations like this. Let us call that range of possible manifestations the power's *manifestation complex*.[7] On the account of powers I'm proposing, no two powers can have the same manifestation complex. That suggests a principle like the following:

> Power_1 = Power_2 if and only if the manifestation complex of Power_1 = the manifestation complex of Power_2.

The left-right conditional should be uncontroversial: a power can't differ from itself, so if $\text{Power}_1 = \text{Power}_2$, then Power_1 and Power_2 must have the same manifestation complex. The grounds for endorsing the right-left conditional need to be spelled out in greater detail, however, since a trope ontology suggests a way of developing

[7] I'm grateful to Joe Vukov for suggesting this term to me.

counterexamples to it: Suppose that *a* is *F* and also that *b* is *F*. Since *a* and *b* are both *F*, it seems reasonable to suppose that the manifestation complex of *a*'s being *F* and the manifestation complex of *b*'s being *F* will be the same: *a* and *b* would manifest themselves in the same ways under all possible circumstances. Hence, the right-hand condition is satisfied. But if properties are tropes, then *a*'s being *F* and *b*'s being *F* are different properties. Tropes are ways that individuals are. *a*'s being *F* is a way that *a* is, whereas *b*'s being *F* is a way that *b* is. Consequently, the left-hand condition is not satisfied. The right-left conditional is therefore false: different powers can have the same manifestation complex. This argument shows that the right-left conditional needs to be treated with care.

The first thing to note here is that *a*'s being *F* and *b*'s being *F* are exactly similar powers. Consequently, they will have exactly similar manifestation complexes. But exact similarity is not yet numerical sameness, and it's evident that there will be differences between the manifestation complex of *a*'s being *F* and the manifestation complex of *b*'s being *F*. Let $M_1[C_i, <a, F>, <c, G>]$ represent the way *a*'s being *F* and *c*'s being *G* would manifest themselves when conjoined under conditions C_i. *b*'s being *F* would have an analogous manifestation if it were conjoined with *c*'s being *G* under C_i, namely $M_1[C_i, <b, F>, <c, G>]$. Even though the manifestation of *a*'s being *F* would be very similar to this manifestation, it would not be the same, for the latter would be a manifestation of *b*'s being *F*, whereas the former would be a manifestation of *a*'s being *F*. The differences in the manifestation complexes of *a*'s being *F* and *b*'s being *F* are traceable to the difference in the powers themselves. Powers are tropes, and even though *a*'s being *F* and *b*'s being *F* are exactly similar, the former is a way that *a* is, and the latter is a way that *b* is. Moreover, since tropes are nontransferable (Section 3.2), the former is essentially a way that *a* is, and the latter is essentially a way that *b* is. Consequently, the manifestation complex for *a*'s being *F* comprises conditions under which *a*'s being *F* manifests itself, whereas the manifestation complex of *b*'s being *F* comprises conditions under which *b*'s being *F* manifests itself. Descriptions of the former will include '*a*' at every place where the manifestation complex of the latter includes '*b*.' As a result, the right-hand condition will not be satisfied unless *a*'s being *F* is identical to *b*'s being *F*, that is, unless $a = b$.

Suppose, then, that *a*'s being *F* and *b*'s being *F* are exactly similar powers, and that $a = b$. In that case, it seems plausible to suppose that *a*'s being *F* would be identical to *b*'s being *F*. This follows if we assume, plausibly, that an individual cannot have two exactly similar powers. Why endorse this assumption? For one thing, denying it would have awkward epistemological consequences similar to those attending categoricalist theories that claim the same property can play multiple causal roles (Section 5.2). Imagine that *a*'s being *F* and *a*'s being F^* are exactly similar powers even though $F \neq F^*$. Because *F* and F^* are exactly similar, it is not possible to tell on any given occasion which power is responsible for a given manifestation. Moreover, if *F* and F^* are exactly similar, then they will have exactly similar manifestation complexes. Because it's in the nature of powers to be ready to go, all that it takes for a

power to manifest itself is for it to be conjoined with reciprocal disposition partners in the right circumstances. Since F and F^* are exactly similar, one of them will be conjoined with reciprocal disposition partners in the right circumstances if and only if the other is. Consequently, both will manifest themselves if either does. As a result, every manifestation of F will be overdetermined by a manifestation of F^* and vice versa—an awkward result. It seems plausible, then, to suppose that an individual cannot have two exactly similar powers. Given this assumption, it follows that if a's being F and b's being F are exactly similar powers with exactly similar manifestation complexes, and $a = b$, then a's being $F = b$'s being F. It follows, in other words, that the right-left conditional is true.

The view of powers I've described implies that a single power can have many different manifestations. Jonathan Lowe (2010a) has argued that this cannot be the case. Powers are identified by their manifestations, he argues, and there can be only one manifestation per power; denying this presents powers theorists with a dilemma:

> We can pose a dilemma for those who suppose that a single power could have more than one manifestation-type. Either those supposedly different types fall under a single unified description or they do not. If they do, then there is really only one manifestation-type. If they don't, then what reason is there to suppose that there is really just *one* power involved rather than two or more—one for each genuinely different manifestation-type?... Once we allow that powers may genuinely have multiple manifestation-types which don't fall under any unified description, it becomes unclear why we should think that a single object may have many different powers rather than just *one*—a power to do all the things it can do. And that would render the notion of power a rather feeble and trivial one. (2010a: 11–12)

There are a few things to say in response to this argument. Let's begin with the first horn of Lowe's dilemma. Consider the example of a batter hitting a baseball 400 feet and his hitting the same baseball 399 feet under slightly different circumstances. It's possible to formulate many descriptions under which both manifestations of the batter's powers fall. In both cases, for instance, it would be true to say that the batter hits a baseball. According to the first horn of Lowe's dilemma, it should follow from this that there is really only one type of manifestation here: hitting a baseball. But if properties are sparse (Chapter 2), then this conclusion does not follow, for in that case, predicates do not correspond one-one to properties, so different descriptions needn't correspond to different types of manifestations. Moreover, it's likely that many of the descriptions under which diverse manifestations fall will be framed in terms of determinable predicates such as 'can hit a baseball' or 'can hit a baseball over 300 feet.' In Section 3.3, I outlined an argument to the effect that determinable predicates do not express any properties distinct from the fully determinate properties that satisfy their definitions. If the batter satisfies the predicate 'hits a baseball' on account of hitting a baseball 400 feet, then in that case, the predicate 'hits a baseball' expresses the fully determinate property of hitting a baseball 400 feet. Conversely, if the batter satisfies the predicate 'hits a baseball' on account of hitting a baseball 399

feet, then in that case, the predicate expresses the fully determinate property of hitting a baseball 399 feet. Consequently, even if it is possible to formulate a single unified description under which two different manifestations of a power fall, it doesn't follow that there is only one type of manifestation.

Consider now the second horn of Lowe's dilemma. Does the postulation of a single power, the batter's power to hit a baseball, threaten to collapse all of his powers into a single power, one that manifests itself not just in his hitting a baseball this or that distance, but in his thinking, feeling, walking, sleeping, and so on? I think not, although the reasons for this will not become fully apparent until Chapter 8. There I argue in detail that an individual's powers are essentially embodied in its parts. This essential embodiment both supports the idea that one power can manifest itself differently under different conditions and provides resources for resisting the collapse of powers that Lowe envisions. Consider these points in order.

According to the hylomorphic theory I defend, an individual's activities are coordinated manifestations of the powers of its parts (Sections 8.1–8.2). When the batter swings the bat, he coordinates the way his parts manifest their powers, and it is that coordination which unifies the diverse physiological events involving his muscles and nerves into a single act of swinging the bat. The batter's ability to impose that coordination on his parts is his power to swing a bat, which is also his power to hit a baseball 400 feet under one set of conditions and 399 feet under slightly different conditions. In both cases he manifests the same power because in both cases he imposes an exactly similar order on the way his parts manifest their powers. The difference between the batter hitting a baseball 400 feet and the same batter hitting a bigger, heavier softball only 300 feet is not a difference in the batter or his powers, but in the size and weight of the ball, the mutual disposition partner with which he interacts. Similarly, it seems likely that the difference between the batter's hitting a baseball 399 feet in one situation and the batter's hitting a baseball 400 feet in a different situation does not depend on a difference in the batter, but in the external conditions in which the ball gets hit: the part of the bat it strikes, the precise angle at which it strikes it, the altitude of the field, the amount of moisture in the air, and so on. The framework of embodied powers thus provides a basis for understanding how a single power can manifest itself differently in different circumstances. That framework also provides resources for avoiding the collapse of powers that Lowe envisions.

I argue in detail in Chapter 8 that all of an individual's powers are essentially embodied in its parts, a claim I call the *embodiment thesis* (Sections 8.2–8.3). If an individual had only a single power, we would expect that individual to have only a single part. Structured individuals like us, however, are composed of many different parts. The batter's power to digest a large meal is embodied in parts different from those which embody his power to swing a bat. Given a hylomorphic framework, the claim that the batter has only a single power which manifests itself in his many diverse activities would have several awkward implications. It would imply, for instance, that the batter has numerous parts which embody none of his powers.

Quite independent of the hylomorphic framework, moreover, it's worth noting that the kind of view Lowe favors, according to which each power has only a single type of manifestation, leads to an awkward and unnecessary proliferation of powers. If we accept Lowe's view, the batter's hitting a baseball 400 feet corresponds to a power different from his hitting a baseball 400.*N* feet, for any real number *N*. Likewise, the batter's hitting a baseball 400 feet along a specific trajectory corresponds to a power different from his hitting a baseball along a trajectory *N* fractions of an inch to the left or right. By contrast with Lowe's view, it seems preferable to posit a single power with different manifestations that vary as a function of external circumstances. Hylomorphists thus have a way of responding to Lowe's dilemma.

4.4 Nomological and Metaphysical Necessity

The identity theory of powers claims that properties are identical to powers. The diamond's hardness is identical to its power to scratch glass. That hardness is the very feature of the diamond that is responsible for the diamond's scratching the mirror when it is raked across the mirror's surface. It is also the very feature of the diamond that is responsible for the diamond's scratching a piece of jade. These powers might appear diverse since they involve different kinds of materials (glass versus jade), yet it does not follow from this difference of materials that the power to scratch glass is different from the power to scratch jade. By analogy, a baseball player is not manifesting a different power when he hits a baseball 400 feet, when he hits a baseball 399 feet under slightly different circumstances, and when he hits a bigger, heavier softball only 300 feet. One and the same power can be manifested in different ways under different circumstances and in conjunction with different disposition partners. The diamond's hardness, that one power, manifests itself when the diamond scratches different materials, both when it scratches glass and when it scratches jade.

The identity of powers and properties can be obscured by the talk I've employed hitherto of properties conferring powers—as if the powers were items separate from the properties that confer them. This way of interpreting the verb 'confer' is at home in, say, political contexts such as when we say, "The Senate conferred on him the power to investigate the allegations." In these cases, the powers are indeed different from the individuals that confer them. Moreover, the individuals confer whatever powers they do only contingently. But these connotations do not obtain in the case of properties and powers, according to the identity theory. Confusion is abetted by talk of powers (plural) conferred by a property (singular). But according to the identity theory, to speak of the variety of powers that one property confers is to speak implicitly of the variety of ways that one property can manifest itself in conjunction with different disposition partners. Just as we can abstract from the property's role as an actualization and focus only on its role as a power, so too we can abstract from its

role as a power for this kind of manifestation and focus only on its role as a power for some other kind of manifestation.

One implication of the identity theory is that properties cannot confer powers other than those they confer in fact. It is not metaphysically possible for the diamond's hardness not to confer on the diamond the power to scratch glass, for according to the identity theory, the diamond's hardness is identical to the diamond's power to scratch glass. It is metaphysically possible for diamonds not to exist in a world, and so it is metaphysically possible for a diamond's hardness not to exist in a world. But it is metaphysically necessary that if a diamond's hardness exists in a world, then that hardness gives the diamond the power to scratch glass. Another way of stating this implication is to say that according to the identity theory of powers, laws of nature are metaphysically necessary; natural necessity is a species of metaphysical necessity (Shoemaker 1980; Swoyer 1982; Ellis 1999, 2001, 2002; Heil 2003, 2005; Molnar 2003). If it is a law of nature in the actual world that diamonds are hard enough to scratch glass, then it is a law of nature in every possible world that diamonds are hard enough to scratch glass.

One objection to the identity theory of powers argues that we can easily conceive of worlds in which diamonds fail to scratch glass, and that conceivability is an indicator of possibility. Consequently, there are possible worlds in which diamonds do not have the power to scratch glass. In that case, however, it is not metaphysically necessary that a diamond's hardness should confer on it the power to scratch glass. The connection between properties and powers is not necessary, therefore, but contingent. Consequently, the argument concludes, the identity theory of powers must be false.

There are at least two ways identity theorists can respond to this argument. Either they can deny that worlds in which diamonds lack the power to scratch glass are genuinely conceivable, or else they can deny that the kind of conceivability we achieve in these cases is a reliable guide to possibility. When it comes to the former strategy, the burden for identity theorists is to explain why we appear capable of conceiving diamonds that lack the power to scratch glass if such diamonds are inconceivable in fact. There are at least two ways that identity theorists might try to explain this (Heil 2003: 93–4). First, they could claim that when we take ourselves to be conceiving of diamonds that lack the power to scratch glass, we are not really conceiving of diamonds at all but of different objects, ones that superficially resemble diamonds but that have altogether different properties and hence lack the power to scratch glass. Second, they could claim that when we take ourselves to be conceiving of diamonds that lack the power to scratch glass, we are in fact conceiving of diamonds, but we are conceiving of them in circumstances in which something is inhibiting their powers—a lubricant on the glass's surface, say.

Critics might nevertheless persist: "We can conceive of this very diamond with its tetrahedrally-arranged carbon atoms coming forcefully into contact with the atoms on the surface of this very mirror without any change in the surrounding

circumstances and yet failing to displace any of the atoms on the mirror." Here identity theorists have the option of shifting to the second strategy: they can accept that the critics are in some sense conceiving what they take themselves to be, but deny that conceiving in this sense is any guide to what is genuinely possible. This would be the case if, say, critics were tacitly employing an epistemic notion of conceivability (Shoemaker 1980: 231). Perhaps when critics say in this context that they are conceiving of diamonds failing to scratch glass, all they are managing to assert is that this situation is not inconsistent with all they know about diamonds. But clearly this kind of conceivability is not necessarily an indicator of possibility, for among other things, critics might not know very much about diamonds. Perhaps if they knew the precise magnitudes of the bonds holding the diamond's carbon atoms in place, and the precise magnitudes of the bonds holding the mirror's atoms in place, they would no longer be able to conceive of the diamond's atoms failing to displace the mirror's atoms. If they insisted that the situation was still conceivable, identity theorists could respond that critics were—perhaps unknown to themselves—imagining a situation in which hidden inhibitory factors were operative or in which something with different properties had been substituted for the diamond or the mirror.

Someone who is deeply committed to contingent laws of nature is not likely to find these strategies very convincing. In Chapter 5, however, I will consider the kind of metaphysic that underwrites the commitment to contingent laws of nature, along with some reasons to reject it. The important point to make here is simply that identity theorists have ways of responding to intuitions that laws of nature are contingent.

4.5 The Identification of Powers with Their Causal Bases

The identity theory of powers also implies that powers are identical to what philosophers like Prior et al. (1982: 251) call their 'causal bases.' There are at least two ways of defining a power's causal basis. The first says that the causal basis of power *P* is a property (or complex of properties) that under the right antecedent conditions is sufficient to cause the manifestation of *P*. The second definition adds two further conditions: first, powers are higher-order properties defined by their causal roles, and second, the causal basis of a power is the first-order property (or complex of properties) that satisfies that role in a particular case in which *P* is attributed to an object. Since the second definition entails the first, Prior et al. claim that there are causal bases in both senses. Identity theorists can accept that there are causal bases in the first sense, but reject the claim that there are causal bases in the second sense. The reason is that they reject the claim that there are higher-order properties (Section 5.3), yet accept that there are powers. If there are no higher-order properties

but there are powers, then powers cannot be higher-order properties, and if powers cannot be higher-order properties, then there cannot be any causal bases for those powers in the second sense.

Despite this disagreement, identity theorists and Prior et al. can agree that the causal basis of, say, the diamond's hardness is its tetrahedral arrangement of carbon atoms. Their disagreement concerns how the diamond's hardness and its tetrahedral arrangement of carbon atoms are related. According to identity theorists, they are identical: powers are in general identical to their causal bases (in the first sense of that term). Prior et al. disagree; according to them, powers are higher-order properties that are realized by the causal bases.

One argument in favor of the identity claim is analogous to the causal role arguments endorsed by psychophysical identity theorists such as David Lewis (1966, 1972).[8] Diamonds can scratch glass because they are hard. This explanation provides an implicit definition of the term 'hardness.' When applied to a diamond, the term refers to the property of the diamond that confers on it the power to scratch glass. What is that property? When scientists study what it is that enables the diamond to scratch glass, they discover that it has that power because it has a tetrahedral arrangement of carbon atoms; they discover, in other words, that having carbon atoms arranged tetrahedrally is what confers on the diamond the power to scratch glass. But if the diamond's hardness is as we've defined it, if it is the (one and only) property that confers on the diamond the power to scratch glass, then it follows by the transitivity of identity that the diamond's hardness must be identical to its having carbon atoms arranged tetrahedrally. More schematically:

(1) The diamond's hardness = the property that confers the power to scratch glass. [Premise]
(2) The property that confers the power to scratch glass = the diamond's having a tetrahedral arrangement of carbon atoms. [Premise]

Therefore, the diamond's hardness = the diamond's having a tetrahedral arrangement of carbon atoms. [From (1) and (2) by the transitivity of identity]

Premise (1) is introduced by definition, Premise (2) is defended empirically, and the conclusion follows validly from the premises. We'll consider two kinds of challenges to this argument. In the remainder of this section, I'll consider arguments that target the conclusion. In Chapter 5, I'll consider arguments that target Premise (1).

Prior et al. (1982: 253–4) have advanced three arguments against the conclusion. The first is a multiple realizability argument similar to those advanced against the

[8] Mumford (1998: 144–54) advances a version of this argument.

psychophysical identity theory in the philosophy of mind.[9] It is possible, say Prior et al., for a single power such as fragility to have different causal bases. Fragility could have molecular bonding α as its causal basis in one individual and crystalline structure β as its causal basis in another. If fragility were identical to molecular bonding α, however, it would be impossible for it to occur without α, and the same is true mutatis mutandis for crystalline structure β. Since it is possible for fragility to occur without α and also without β, fragility cannot be identical to either. Since the same is likely true of other powers, we can conclude that powers and their causal bases must be distinct.

There are at least three ways that defenders of the identity theory of powers can respond to this argument (Figure 4.1). First, they can postulate narrowly individuated powers. Second, they can postulate broadly individuated causal bases, and third, they can postulate a combination of the two.[10] Let us consider these responses in order.

The multiple realizability argument suggests that powers and causal bases are correlated one-many (Figure 4.1, column A). A single power P is correlated with multiple causal bases $B_1, B_2, \ldots, B_n$. Suppose, however, that powers are individuated more narrowly than Prior et al. suppose. Instead of there being a single power P, there are many narrowly individuated powers $P_1, P_2, \ldots, P_n$, each of which corresponds to a single causal basis (Figure 4.1, column B).[11] In that case, the multiple realizability argument fails to show that powers are not identical to their causal bases, for it is possible for the many narrowly individuated powers to be identical to the diverse causal bases: $P_1 = B_1, P_2 = B_2, \ldots, P_n = B_n$. The term 'fragility,' for instance, might not refer to a single kind of power found in porcelain vases and glass windows. It might be instead an imprecise term that refers to one kind of power in porcelain and

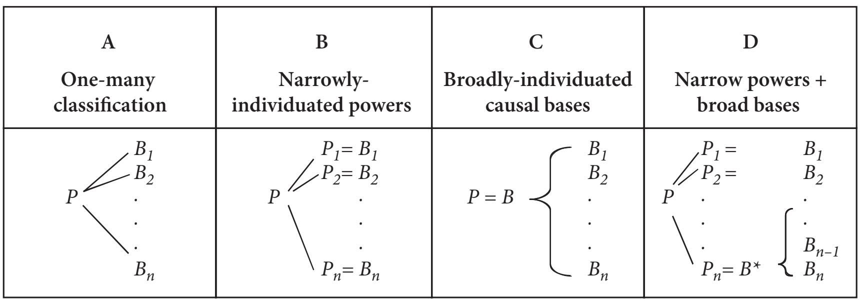


Figure 4.1 Responses to the multiple realizability argument.

[9] For a discussion of the argument, see Jaworski (2011: chapter 6), and also the entry 'Mind and Multiple Realizability' in the Internet Encyclopedia of Philosophy, <http://www.iep.utm.edu/mult-rea/>.

[10] These responses are analogous to the ones that psychophysical identity theorists can make in response to multiple realizability arguments in philosophy of mind (Jaworski 2011: 134–6). See also Jaworski's entry 'Mind and Multiple Realizability' in the Internet Encyclopedia of Philosophy, <http://www.iep.utm.edu/mult-rea/>.

[11] This is Heil's (2003: 116) preferred strategy. Psillos (2006: 149) describes a similar idea.

another kind of power in glass. By analogy, the term 'jade' was originally taken to refer to a single kind of mineral, but it was later discovered that it was actually an imprecise term that referred to at least two different kinds of minerals: jadeite and nephrite. If 'fragility' is analogous, if it does not refer to a single power, but to many different powers, then it is possible for those many powers to be identical to diverse causal bases: fragility-of-porcelain = molecular bonding α, fragility-of-glass = crystalline structure β, and so on. If that is the case, then Prior et al. have failed to show that powers are not identical to their causal bases.

A second way that identity theorists can respond is by postulating broadly individuated causal bases (Figure 4.1, column C). Prior et al. assume that molecular bonding α and crystalline structure β are in fact different kinds of causal bases, but it might turn out that they are not. Bases α and β might have important commonalities that qualify them as instances of a broader kind of causal basis, γ, whose instances are correlated one-one with instances of fragility. Suppose, for instance, that the terms 'molecular bonding α' and 'crystalline structure β' are analogous to 'electricity' and 'magnetism,' terms which were originally taken to refer to different phenomena, but which were later revealed to be manifestations of a single overarching type of phenomenon (a discovery still reflected in the name 'electromagnetism'). In this case too, the multiple realizability argument fails to show that powers are not identical to their causal bases, for it fails to show that fragility is not identical to physical state γ.

Finally, identity theorists can combine the two aforementioned strategies. They can claim that our classification of powers and causal bases can both be altered in ways that yield one-one power-basis correlations (Figure 4.1, column D). In the absence of compelling reasons to endorse Prior et al.'s way of individuating powers and causal bases, we can conclude that the multiple realizability argument fails to show that the identity theory of powers is false.[12]

Prior et al.'s multiple realizability argument claims that it is possible for a power to be correlated with different causal bases. Their second argument claims that it is possible for a causal basis to be correlated with different powers. It is possible, for instance, for the same causal basis to confer a power in one case and not to confer it in another. This could happen, they say, if an object had additional properties that 'swamped' the effects of its causal basis. A given vase might have molecular bonding α, together with some other property, S, that prevents α from being effective. If that were the case, then every fragile object might have α, but it would not follow that every object with α would be fragile since objects that had S in addition to α would not be fragile.

McKitrick (2003: 360–1) argues that this second argument fails. The reason is that Prior et al.'s example seems to be simply a case of masking (Johnston 1992: 233;

[12] McKitrick (2003: 357–61) shows that the argument fails for different reasons.

Molnar 2003: 130). Masking occurs when something has a power but something else prevents that power from manifesting itself. An example would be a fragile vase that has been stuffed with packing material which prevents the vase from breaking. Intuitively, it seems that the vase is still fragile; the packing material simply prevents the fragility from manifesting itself. Prior et al. provide no reason to think that the imagined property *S* does not do the same. In that case, however, there is no reason to conclude that it is possible for something to have α without being fragile, and so there is no reason to conclude that fragility and α must be distinct.

Finally, Prior et al. argue that if powers were identical to their causal bases, then necessarily they would be correlated with those bases one-one, and yet they are not. In support of the argument's major premise, Prior et al. argue that property names such as 'fragility' and 'molecular bonding α' are rigid designators. Consequently, if it is true that fragility = molecular bonding α in the actual world, then it is true in every possible world. It is nevertheless possible, they say, that fragility might be correlated with different causal bases in different possible worlds—with, say, molecular bonding α in w_1 and with crystalline structure β in w_2. As a result, fragility cannot be identical to α in every possible world, but since 'fragility' and 'molecular bonding α' are rigid, it follows that fragility and molecular bonding α cannot be identical in any possible world. Hence, powers and their causal bases must be distinct.

The crucial premise of this argument is that powers and their causal bases are correlated only contingently. Prior et al. defend this premise by appeal to two further claims. First, powers are defined by counterfactual conditionals. Fragility, for instance, might be defined as the power such that if an object having this power were dropped, it would break. Second, it is possible, they say, for different causal bases to satisfy these definitions. This second premise, they say, is undisputed. Yet, we have already considered a view that disputes it, namely the view that postulates narrowly individuated powers. According to that view, molecular bonding α and crystalline structure β are not causal bases for one and the same power; the former is, rather, the causal basis of fragility-in-porcelain, the latter of fragility-in-glass. In addition, there are likely to be differences between α and β. If there are not, then there is no reason to accept Prior et al.'s assumption that α and β are distinct, and without that assumption, the examples of α and β fail to support the premise that fragility could have different causal bases. Prior et al.'s argument would fall prey to the postulation of broadly individuated causal bases discussed a moment ago. If there are no causally relevant differences between α and β, then there is no good reason to think that α and β are not just instances of a broader type of causal basis, γ. Suppose, however, that α and β are in fact different. Suppose, for instance, that objects with α are slightly more resistant to breaking under impact than objects with β. In that case, there is good reason to think that α and β are bases for different powers: the power to break when subjected to a force of *N*, say, versus the power to break when subjected to a force of N^*. Again, if there were no differences of this sort, then there would be no grounds to suppose that α and β are really distinct.

Prior et al. say nothing that rules out the postulation of broadly individuated causal bases, nor do they say anything that rules out the postulation of narrowly individuated powers. But in the absence of good reasons to reject one or the other, there is no good reason to accept their premise that a single power can have many different kinds of causal bases, and consequently, their argument fails to show that powers and their causal bases are distinct.

4.6 Identity Theoretic Solutions to Dispositionalist Problems

The identity theory of powers is closely related to pure dispositionalist theories of powers like Shoemaker's (1980). Like the identity theory, pure dispositionalist theories claim that properties have dispositional essences. Unlike the identity theory, however, pure dispositionalist theories claim that the identity conditions of properties are exhausted by the powers they confer on things. One advantage of identifying powers with their causal bases, as identity theorists do, is that it offers attractive solutions to the problems confronting pure dispositionalist theories. These include the problem of missing bases, the always packing, never traveling objection, and what Mumford (2006) calls the 'problem of Being.'[13] I'll consider these in order.

The problem of missing bases argues that pure dispositionalist theories must be false because they imply that powers have no causal bases (Molnar 2003: chapter 8). Prior et al. (1982: 251–3) argue that every power must have a causal basis in at least the first sense defined in Section 4.4: a condition or collection of conditions that are sufficient for a power to manifest itself. If there were no such conditions, they argue, then there would be no powers, for there would be no manifestations that they would be powers for. Many philosophers have assumed, however, that the causal basis of a power must provide lower-level explanations for the power's manifestations; they have assumed, in other words, that when we explain the manifestation of *a*'s power, we must do so in terms of the properties of *a*'s parts. A manifestation of the diamond's hardness, for instance, must be explained by appeal to the properties of the diamond's carbon atoms. This poses a problem when it comes to the powers of fundamental physical particles,[14] for fundamental physical particles have no parts (Harré and Madden 1975: 155; Martin 1993: 184; Ellis 2001: 114; 2002: 74–5; Molnar 2003: 131–2; Mumford 2006). Consequently, it is impossible to explain the manifestations of the powers of fundamental physical particles in terms of the properties of their parts. A pure dispositionalist view thus fails to account for the powers of fundamental physical particles.

[13] Jacobs (2011) advances a different argument against pure dispositionalist views. He argues that nothing is capable of satisfying the identity conditions for powers that pure dispositionalists provide.

[14] Psillos (2006: 151–4) argues that the claim that the properties of fundamental physical particles are powers is unsupported.

Molnar (2003: 132–42) considers several responses to the problem. They include (1) claiming that the powers of fundamental physical particles have causal bases that are simply unknown, (2) claiming that the causal bases are not microphysical, but global, (3) claiming that fundamental physical particles do not really have powers, and (4) claiming that some powers have causal bases but others do not. He finds each of these responses problematic, but ultimately plumps for a version of (4): many powers have causal bases, he says, but the powers of fundamental physical particles do not. There are several worries about this solution. One is that it looks ad hoc. It appears to introduce a distinction between powers with causal bases and powers without them, not for any principled reason, but simply as a way of responding to the problem. In addition, there are at least two objections to this solution.

The first objection Molnar calls the 'always packing, never traveling' objection.[15] It claims that a view committed to powers without causal bases—'ungrounded powers' henceforth—is a view that lacks resources for explaining the actuality of manifestations (Campbell 1976: 93–4; Foster 1982: 67–72; Robinson 1982: chapter 7; Swinburne 1980; Blackburn 1990; Martin 1997: 213–17; Armstrong 1996c: 96; 1997: 80; 1999: 31–2; 2005: 314). Armstrong states the objection as follows:

> Suppose that the world consists of particulars having properties in the narrow sense and related to each other by external relations . . . Now suppose that these properties and relations are nothing but powers. It will follow that the manifestations of these powers, when they occur, can themselves be nothing but cases of particulars coming to have certain powers. After all, manifestations are nothing but certain particulars coming to have certain properties, and on the theory being criticized all properties dissolve into powers. But could there be a world of this sort? Powers must surely issue in manifestations that are something more than just powers. A world where potency never issued in act, but only in more potency, would be one where one traveled without ever having the possibility of arriving. (2004a: 139)

Molnar (2003: section 11.2) considers several versions of the argument, and finds them wanting. The best version, he thinks, is Martin's:

> Dispositionalists believe that all that appears to be qualitatively intrinsic to things just reduces to capacities/dispositions for the formation of other capacities/dispositions for the formation of other capacities/dispositions for the formation of . . . And, of course, the manifestations of any disposition can only be further dispositions for . . . This image appears absurd even if one is a realist about capacities/dispositions. (1997: 215)

According to Molnar, Martin's version of the objection does not provide a compelling reason to reject ungrounded powers. The reason is that it depends on the assumption that there are intrinsic qualities, and as we saw in Section 4.1, Molnar

[15] It's worth mentioning that even though Molnar rejects a pure dispositionalist view like Shoemaker's (1980), he must still confront the objection when it comes to the powers of fundamental physical particles.

(2003: 178) argues that there is good reason to believe that fundamental physical particles do not have any qualities at all. If Molnar is right, then Martin's argument does not threaten his view.

But things are different with Armstrong's version of the argument quoted above. It does not posit intrinsic qualities, so Molnar's objection to Martin does not apply to it. Armstrong's version of the argument concerns actualities in general. Pure dispositionalist views are reticent about the notion of actuality, the argument claims, and as a result, it remains unclear how these views can account for the actualities (that is, the manifestations) that powers are supposed to bring about. The same is true mutatis mutandis of views that posit ungrounded powers. If there are no causal bases at a fundamental physical level but only powers, then it remains unclear how anything actual can happen at a fundamental physical level. If the exercise of fundamental physical power *P* results in an instance of property *F*, and fundamental physical properties are nothing but powers, then *F* itself is just a power. But if *F* itself is just a power, then the exercise of *P* has resulted simply in bringing about a different distribution of powers. "It is like a promissory note," Martin says, "that may be actual enough but if it is for only *another* promissory note which is [for only *another* promissory note which is] . . . , that is entirely too promissory" (1997: 215). What Molnar says appears to fall short, therefore, of providing a completely satisfactory response to the always packing, never traveling objection.

The always packing, never traveling objection is closely related to another (Mumford 2006: 485; Psillos 2006; Bauer 2012). Mumford calls it 'the problem of Being':

> To be a disposition is just to be directed towards some possible manifestation. To be an ungrounded disposition is to be so directed and nothing else. In particular, it is for there to be no micro-structural ground to the directedness . . . But if such a property is ungrounded, what in the world is it that is directed towards some possible manifestation? Such a property looks like no property at all. It is nothing more than the possibility of some future property when there is a manifestation. An ungrounded disposition has no *Being* between its manifestations. There is no *thing* that is directed towards any other thing, no other states, properties or facts in the world that cause or ground those manifestations—manifestations that need never be actualized. (2006: 483)

Any theory of powers, it seems, must explain what enables a power to persist when it is not manifested. As Bauer puts it:

> [G]iven the modal nature of pure dispositions it seems there must be some explanation of their being when not manifesting . . . [I]f there is no explanation of the being of pure dispositions, then such entities would just seem to 'hang' ontologically on nothing . . . Even if pure dispositions do 'hang' in this way, some explanation of what this amounts to is in order. (2012: 140–1)

The most likely candidate for explaining a power's persistence is the power's causal basis. The persistence of the basis explains the persistence of the power. Yet, if powers are ungrounded, if they have no causal bases, then it is no longer clear what explains their persistence.

The always packing, never traveling objection and the problem of Being pose difficulties for any view that posits ungrounded dispositions, both views that posit them tout court, as Shoemaker's (1980) view does, and views that posit them in a more limited way, as Molnar's (2003) and Ellis' (2002) views do. An attractive feature of the identity theory of powers is that it offers solutions to both problems, as well as to the problem of missing bases.

The identity theory implies a solution to the problem of missing bases that Molnar does not consider: it identifies powers with their bases, and rejects the idea that an explanation of a power's manifestation must be given by appeal to the parts of the individual whose power it is.[16] Once we reject the claim that the explanation of a power's manifestation must appeal to something's parts, then it is no longer problematic for something, such as a fundamental physical particle, to have powers and yet lack parts. This does not imply that the powers of fundamental physical particles have no causal bases; it implies rather that those powers are their own causal bases. Powers are not ungrounded, they are instead self-grounded.[17] One might worry that positing self-grounded powers in this way is an ad hoc maneuver. But this does not appear to be the case. The reason is that the identity theory implies that all powers are grounded in themselves since it implies that every property is identical to its causal basis.[18] There is thus a principled basis for claiming that all powers are self-grounded.

Psillos (2006: 147–8) advances two arguments against the claim that powers can be self-grounded. The first argument is based on the premise that a self-grounded power would have to explain both why the power is manifested in one case and why it is unmanifested in another. The vase's fragility, for instance, would have to provide a basis for explaining both why the vase would break when struck by a hammer and also why it would remain intact when not struck by a hammer. Psillos insists that it is unclear how one and the same power could explain both things, and thus concludes that powers cannot be self-grounded.

[16] McKitrick (2010) and Marmodoro (forthcoming) discuss solutions to the problem different from the one offered here.

[17] McKitrick (2003) and Mumford (2006) use the term 'ungrounded' to refer to a disposition that has no causal basis distinct from itself. Consequently, their use of the term 'ungrounded' is compatible with a disposition having no ground at all, and also with a disposition having itself as a ground. I use the term 'ungrounded' more narrowly to refer to a disposition that has no ground at all, and I've introduced the term 'self-grounded' to refer to dispositions that ground themselves. I think this is clearer than using the term 'ungrounded' in the way McKitrick and Mumford do.

[18] McKitrick (2003) calls properties of this sort 'bare dispositions,' powers that have no distinct causal bases. Bare dispositions do have causal bases, they are just not distinct from them; each bare disposition, in other words, is its own causal basis.

The crucial premise in Psillos' argument is that ungrounded powers must explain both their manifestations and their unmanifestations. Psillos suggests that unless self-grounded powers (what he calls 'bare' powers) do this, there is nothing at all that explains why a particular power is unmanifested:

> The "bare" power is bound to do more than causally contribute to its own manifestation. It also causally contributes to the *absence* of this manifestation . . . For, considering the object that has the "bare" power *in and of itself*, there is nothing else . . . which causally contributes to the *lack* of the manifestation of the power . . . because all there is is the "bare" power. (Psillos 2006: 147)

Psillos' argument is unconvincing. If we examine our actual explanatory practices, there seems little reason to think that we must explain the absence of a manifestation by appeal to the presence of a power for it. Consider two vases, A and B: A breaks, B doesn't. What explains the difference? We do not say, "The difference is that they are both fragile." We say instead, "The difference is that A was struck by a hammer and B not." Psillos claims that if fragility is self-grounded, then this way of explaining the difference between A and B removes from the power any causal role in the breaking. "[T]he causal burden," he says, is instead, "shifted to the external stimulus" (147). But this seems manifestly false. If the vase were not fragile, then clearly it would not break when struck by the hammer. The contribution of the vase's fragility seems clear. It is unclear, moreover, why Psillos thinks the power's self-grounded status should make any difference to this kind of explanation since the postulation of self-grounded powers is not incompatible with the explanatory pattern just described.

Second, Psillos argues that if the same ungrounded power does not explain both why it is manifested in some cases and why it is unmanifested in others, then we end up postulating an embarrassingly large number of powers: a power to explain why vase B remains intact, another to explain why it remains intact when tapped only lightly with the hammer, and so on. But, he says, postulating this many powers offends against Ockham's razor.

Two remarks are in order about this argument. First, Ockham's razor is a comparative principle. We offend against it only if there is a more economical alternative that is in other respects comparable to our favored account. Yet, it is not clear that there is a more economical alternative in this case. Psillos (144) says that he favors a view of powers like Armstrong's (1983, 1997), which postulates categorical properties in conjunction with laws of nature. Clearly, a view like this is going to have to explain why vase B remains intact, as well as why vase B remains intact when tapped only lightly with a hammer. Presumably, it will explain these things by appeal to different laws of nature. If that is the case, however, then this view seems no more economical than the self-grounded powers view, for if the latter has to postulate a different power to account for each of these phenomena, then it seems likely that the former will have to postulate a different law of nature to account for each of them. Perhaps Psillos has some reason to deny this; perhaps he thinks that a number of different phenomena

can all be explained by appeal to a single law. The problem with this strategy is that there is nothing to prevent an exponent of self-grounded powers from making an analogous claim. We have already seen that identity theorists claim that a single power can ground many different manifestations (Sections 4.2 and 4.3). Why, then, should it not be that a single power grounds vase B's remaining intact, vase B's remaining intact when only lightly tapped with a hammer, and so on? For these reasons, the appeal to Ockham's razor does not appear to succeed.

Second, Psillos assumes that the power that explains the aforementioned manifestations must be vase B's fragility. But this seems false in some explanatory contexts, and innocuously true in others. Consider the following examples:

(1) Why did vase B remain intact, in contrast to shattering, given that A shattered, in contrast to having remained intact?
(2) Why did vase B remain intact, in contrast to crumbling spontaneously?
(3) Why did vase B remain intact, in contrast to shattering, given that it was tapped lightly with a hammer?

The correct answer to (1) seems to be that vase B remained intact, in contrast to shattering, because it was *not* struck by a hammer (unlike vase A). Here the correct explanation does not appeal to fragility; in fact, it does not appeal to the presence of a power at all, but to the absence of an appropriate stimulus. The correct answer to (2), on the other hand, seems to be that vase B has a certain molecular structure that resists crumbling, and the same goes mutatis mutandis for (3): vase B has a molecular structure that resists shattering when tapped lightly with a hammer. Is this molecular structure identical to the vase's fragility? There is no reason why identity theorists should resist this claim, for as we've seen, the same power can manifest itself differently in different circumstances. The same power/molecular structure that resists spontaneous crumbling and light tapping cannot resist hard rapping. Resisting crumbling, resisting light tapping with a hammer, and shattering when struck (hard) with a hammer are just different manifestations of the same power, just as hitting a baseball 400 feet, hitting a baseball 399 feet, and hitting a softball 300 feet can be manifestations of the same power to bat. Psillos' claim that this view leads to an embarrassing proliferation of powers seems to be based on the assumption that a single power cannot have different manifestations. Since he does nothing to defend this assumption, we can conclude that his argument does not succeed in ruling out self-grounded powers.

Consider now the always packing, never traveling objection. Because the identity theory makes explicit each property's status as a stable manifestation (or actualization) of something's powers, it is able to sidestep the objection. The diamond's hardness is a stable manifestation of a power the carbon atoms have. The diamond's hardness is itself a power, to be sure, but the identification of this power with the diamond's having a tetrahedral arrangement of carbon atoms ensures that the identity theory does not

imply the mere shifting around of potencies that pure dispositionalist views are accused of implying.

In addition, causal bases have generally been considered good candidates to explain the persistence of powers. Since the identity theory identifies powers with their causal bases, there is good reason to think that the identity theory provides a solution to the problem of Being as well.[19] What explains the persistence of the diamond's hardness when that power is not manifesting itself in, say, the scratching of glass? The answer is that the power persists as an actual property of the diamond, a property we express using the predicate 'has a tetrahedral arrangement of carbon atoms.' We also express this property using the predicate 'is hard,' but in doing so, we bring out a different theoretical role the property plays, not its role as a stable manifestation, but its role as a power directed toward other manifestations.

It's noteworthy that some dispositionalists respond to the problem of Being in a similar way. Stephen Mumford is an example:

> [This objection] assumes that a power is a potential only and not at all actual in its own right . . . [W]e are presented with a very puzzling interpretation of powers. If the power is not actual unless its manifestation is, then in what sense is it a power to F? In what sense is it anything at all? The danger of this view is that it treats powers as nothing more than mere potentialities but thereby ignores the obvious point that to be potent (as opposed to potential) is to be actual . . . Potent means powerful, which is something very different from being potential, meaning not yet actual. Those who favor powers regard them as potent rather than potential. Hence . . . they are also things in their own right. (Mumford 2004: 174; cf. Mumford 2006: 485)

Mumford's response to the problem of Being stops short of identifying powers with their causal bases (a view defended in Mumford 1998), but his emphasis on the stable actuality of powers does seem to bring his pure dispositionalist view a step closer to the identity theory.

For our part, we've seen that there are good reasons for identifying powers with their causal bases. If there is a remaining reason to hedge, it would seem to concern the powers of fundamental physical particles. Critics might object that the solution I've sketched to the problem of Being is no solution at all—at least not when it comes to the powers of fundamental physical particles. In cases like the diamond's hardness, the concept of the causal basis of a power is independent of the concept of the power itself. The concept of a tetrahedral arrangement of carbon atoms, for instance, is independent of the concept of hardness. In the case of fundamental physical particles, however, there is no concept of a causal basis that is independent of the concept of the particles' powers themselves.

[19] Psillos (2006: 151–2) in fact claims that the Martin-Heil view provides one way of solving the problem of Being.

In response, identity theorists can grant the premise that there is no concept of a causal basis for the powers of fundamental physical particles that is independent of the concept of the powers themselves. What they deny is the argument's implicit assumption that in order for powers to be grounded in their causal bases, the powers and their bases must answer to different concepts. Their approach to the problem implies that this assumption is false. In the absence of a compelling argument to the contrary, there is no reason for them to abandon this position, and hence no reason for them to abandon their commitment to powers being identical to their causal bases—including the powers of fundamental physical particles. These powers are not groundless, they say: they are grounded in themselves—all powers are. If there is a difference between the powers of fundamental physical particles and the powers of other things, it is that the latter powers, but not the former, answer to multiple independent concepts. What explains this difference is not altogether clear. Perhaps it is simply that the powers of fundamental physical particles are not readily experienced by us. Whatever the explanation, it is largely an epistemological one, according to identity theorists: it concerns our ways of conceiving things, not the things themselves.

A final objection to pure dispositionalist theories like Shoemaker's (1980) is advanced by John Hawthorne (2001). Hawthorne calls Shoemaker's view 'causal structuralism.' According to causal structuralism, as Hawthorne understands it, every natural property has a causal profile, a set of powers that it confers on its possessors, and a property's causal profile is both sufficient and necessary for being that property. Hawthorne's objection to causal structuralism concerns the possibility that two different properties might have symmetrical causal profiles in a given possible world. Suppose, for instance, that A, B, C, and D are properties that stand in the following relations in world *w*: A causally necessitates C, B causally necessitates C, and the conjunction of A and B causally necessitates D. One way of defining properties is to construct a Ramsey sentence for their causal profile. Ramsey sentences replace the names of the properties with variables, and prefix quantifiers corresponding to each of the variables. A Ramsey sentence defining property A, for instance, would be '$\exists x \exists y \exists z \exists v$ (x causally necessitates z, y causally necessitates z, and the conjunction of x and y causally necessitates v).' The problem for causal structuralism is that this sentence can be used to define not only property A, but also property B. As a result, A and B cannot be individuated by appeal to their causal profiles in *w*. Since causal structuralism is committed to a property's causal profile being sufficient to distinguish it from other properties, it looks as though causal structuralism must be false.

There seems to be a way of developing Hawthorne's argument into an objection to the identity theory of powers. The identity theory, as I've described it, denies that properties are individuated by the causal profiles they have in any given possible world. Rather, what distinguishes one property from another is its manifestation complex, the full range of its possible manifestations (Section 4.3). No two properties

can manifest themselves the same way in all possible circumstances. Consequently, even if A and B were to have symmetrical profiles in world *w*, there would be another possible world in which their causal profiles differed. The identity theory implies, however, that a property's causal profile is essential to it; it therefore implies that a property's causal profile cannot differ from one world to another. So if A and B have different profiles in some possible world, it follows that they must have different profiles in all possible worlds. This implies that the kind of world Hawthorne envisions must be impossible: there can be no world in which two properties have symmetrical causal profiles. But, says Hawthorne, it seems "intuitively possible" that two properties could have symmetrical causal profiles. In addition, that possibility is implied by a combinatorialist account of modality, according to which any conjunction of causal profiles yields a logically consistent Ramsey sentence corresponding to a possible property.

In response, I think an identity theorist has no choice but to reject Hawthorne's intuitions as well as the combinatorialist account of modality he describes. When it comes to the former task, I have already discussed some strategies for defusing intuitions that threaten the identity theory's implications (Section 4.4). When it comes to the latter task, the identity theorist seems committed to constraining the range of possible worlds in a way that Hawthorne's envisioned combinatorialism does not. Some are likely to see this as a significant cost for endorsing the identity theory. On balance, though, I think that the benefits of the identity theory make the cost worthwhile.

4.7 Objections to the Identity Theory of Powers

We have considered several objections to the identity theory of powers. They include the objection that the identity theory is committed to Meinongian nonactual entities (Section 4.2), the objection that properties cannot be identical to powers since the laws of nature could have been different from what they are (Section 4.4), and several objections to the claim that powers are identical to their causal bases (Section 4.5). We've seen that identity theorists have responses to each of these objections. In this section, I consider a final objection to the identity theory advanced by David Armstrong. It claims that the identity theory makes it unclear how a property's dispositional roles are related to its nondispositional roles (Armstrong 1996c: 96–7; 2005: 314–15).

The relations between a property's dispositional and nondispositional roles must either be necessary or contingent, yet neither seems satisfactory, says Armstrong. Suppose, on the one hand, that the relation is contingent. In that case, it would be possible for, say, the diamond's hardness, conceived nondispositionally, to be correlated with different dispositions such as the disposition *not* to scratch a mirror if raked across its surface. Yet, this is at odds with the identity theory, which claims that the diamond's hardness is identical to its power to scratch glass. The relation between

the hardness and the power to scratch glass must not be contingent, therefore, but necessary. But if the relations between a property's dispositional and nondispositional roles are necessary, then the view becomes obscure, says Armstrong, for something must explain why the roles are related necessarily, and it is difficult to see what could explain this. It seems that defenders of necessary relations must posit those relations as brute unexplainable matters of fact, and this seems ad hoc. Armstrong concedes that identity theorists have a response: the reason why the diamond's hardness is necessarily correlated with the diamond's power to scratch glass is that the diamond's hardness is identical to the diamond's power to scratch glass. Armstrong meets this response with an incredulous stare:

> I confess that I find this totally incredible. If anything is a category mistake, it is a category mistake to identify a quality—a categorical property—and a power, essentially something that points to a certain effect. They are just different, that's all. (Armstrong 2005: 315)

There are several things identity theorists can say in response. First, an incredulous stare is not by itself an argument—especially when the intuitions that motivate the stare are not very widespread—certainly not as widespread as, say, the intuitions that oppose David Lewis' modal realism. Someone might insist that Armstrong's objection is more than an incredulous stare, that Armstrong's crucial premise is that identifying qualities with powers is a category mistake. If this is the case, however, then it's difficult to see how Armstrong's objection avoids begging the question against identity theorists. To assume at the outset that qualities and powers are "just different," as he says, is simply to assume at the outset that the identity theory is false. It doesn't prove the identity theory is false, it takes its falsity as a given.

The foregoing points by themselves are enough to dispense with Armstrong's objection, for if there are no substantial reasons to reject the identity theory, then there are no substantial reasons not to grasp the second horn of Armstrong's dilemma. There are nevertheless a few further points worth noting. First, Armstrong's intuitions on this point seem to be heavily theory-laden. His insistence that categorical properties and powers are "just different" looks merely like a byproduct of his own view of categorical properties and powers. We consider this view in greater detail in Chapter 5, along with some reasons for rejecting it. The important point here is that if Armstrong's intuitions are based on his theory, then the evidential value of those intuitions is no better than the reasons for or against that theory. Moreover, as I've already indicated, insofar as Armstrong's theory and the identity theory of powers are competitors, Armstrong's intuitions beg the question against the identity theory: they assume rather than prove that the identity theory is false. Finally, there is something ironic about Armstrong's intuitions, given his other philosophical commitments. In the philosophy of mind, for instance, Armstrong (1968) is a psychophysical identity theorist: he looks to identify mental and physical properties—properties that many philosophers of mind would insist are "just different." Why Armstrong's intuitions should deem one kind of identification plausible

and an analogous identification incredible seems at least peculiar, if not ad hoc. Perhaps Armstrong's incredulity is motivated by Martin and Heil's talk of qualities. Since talk of qualities suggests characteristics that are experienced, Armstrong might take the identity theory of powers to imply that all properties are capable of being experienced. We saw in Section 4.1, however, that the identity theory does not imply this. On the basis of the foregoing considerations, I conclude that Armstrong's argument against the identity theory of powers does not succeed.

4.8 Conclusion

I've articulated an identity theory of powers like the one defended by C. B. Martin and John Heil. On this view, one and the same property plays a variety of theoretical roles which we express using different vocabularies. Sometimes we use a dispositional vocabulary, other times we use a nondispositional one. These different vocabularies create the impression that there are different kinds of properties: dispositional and categorical (or qualitative). According to the identity theory, though, these vocabularies describe the very same properties; they just bring out the different theoretical roles these properties play. It is thus a mistake, according to this view, to distinguish between categorical and dispositional properties. At best, the categorical-dispositional distinction corresponds to a difference among predicates, not properties. Properties, rather, are one and all powers that are essentially directed toward their manifestations. This directedness is a primitive feature of theirs. Every power is "ready to go," as Martin puts it; it manifests itself when conjoined with mutual disposition partners under the right circumstances. According to the view of powers I've defended, powers are individuated by their manifestation complexes, the range of possible ways in which they can manifest themselves in conjunction with different disposition partners. No two powers can manifest themselves the same way in all possible circumstances. Consequently, $\text{Power}_1 = \text{Power}_2$ if and only if the manifestation complex of Power_1 = the manifestation complex of Power_2.

I've considered some reasons for accepting the identity theory of powers, and have argued that it enjoys certain advantages over competing dispositionalist theories (Section 4.5). I've also argued that identity theorists have responses to the objections that have been advanced against their view (Sections 4.2–4.4 and 4.6). In Chapter 5, I'll round out the case for the identity theory by considering its principal competitors and some of the reasons for rejecting them. Those competitors include, first and foremost, categoricalist theories of powers such as those defended by David Armstrong and David Lewis.

5

Competing Theories of Powers

5.1 Conditional Analyses

The identity theory of powers is at odds with several others that have appeared in the literature. I have already explained some of the advantages it enjoys over pure dispositionalist theories (Section 4.6). In this section and the next, I consider some of the reasons to favor it over categoricalist theories. Categoricalists deny that properties have dispositional essences. Among other things, this implies the falsity of one of the identity theory's central claims, namely the claim that the diamond's hardness is what confers on the diamond the power to scratch glass. Some categoricalists argue that what confers this power is not the diamond's hardness, but a contingent law of nature (Armstrong 1983, 1997, 2010). Others argue that what confers the power is the causal basis of the hardness, which is distinct from the hardness itself (Prior et al. 1982). In the sections that follow, I will consider these views in greater detail. I cannot offer a comprehensive treatment of the arguments for and against them since that would take us well beyond the scope of this discussion. I nevertheless hope to make clear my own reasons for finding the identity theory of powers preferable.

Before turning to categoricalist theories of properties, it is helpful to note that the identity theory is at odds with accounts of powers that look to provide reductive analyses of disposition ascriptions. Disposition ascriptions are statements such as 'Table salt is water soluble,' which we often use to express something's possession of a power. Disposition ascriptions support counterfactual conditionals such as 'If this table salt were placed in water, then it would dissolve.' Philosophers such as Ryle (1949), Quine (1960), and Lewis (1997) have attempted to analyze disposition ascriptions in terms of the corresponding counterfactual conditionals. The problems these attempts have faced are at least twofold. First, whatever else might be involved in analyzing disposition ascriptions in terms of counterfactual conditionals, this position implies that the truth of the conditionals is sufficient and necessary for the truth of the ascriptions. Yet, there appear to be counterexamples both to the claim that the conditionals are sufficient for the ascriptions (Johnston 1992; Bird 2007) and to the claim that they are necessary (Martin 1994; Bird 1998, 2007). Second, irrespective of whether the conditionals are sufficient or necessary, it remains unclear what it is that makes the conditionals true, if not the properties of the actual things they are true of. What makes it true, for instance, that this salt would dissolve if

placed in water? It is difficult to see what it could be other than actual properties of the salt and the water, and to that extent it is difficult to see how exponents of conditional analyses could answer this request for an explanation other than by positing the truth of the counterfactual conditional as a brute matter of fact. By contrast, the account of properties and powers I've described has an explanation ready to hand: what makes it true that this salt would dissolve if placed in water is that salt and water have the power to interact in this way.

5.2 Categoricalist Theories: Quidditism plus Laws of Nature

Consider now views that posit purely categorical properties, properties that are completely nondispositional. Armstrong's (1996a, b, c, 1997, 2004a, 2005) and Lewis' (1986a) theories of properties are examples. Both deny that properties have dispositional essences; the essences of properties are instead what Robert Black (2000) has called 'quiddities.' Quiddities are supposed to be primitive principles of identity and distinctness for properties which are analogous to haecceities, primitive principles of identity and distinctness for individuals. According to exponents of haecceities, individuals *a* and *b* are distinct if and only if they have different haecceities. Whatever possible similarities *a* and *b* might have, their possession of different haecceities ensures that they are different individuals, and conversely, *a*'s and *b*'s possession of the same haecceity ensures that *a* and *b* are numerically the same individual irrespective of whatever possible differences they might have. Quiddities are supposed to do for properties what haecceities do for individuals. According to quidditists, property *F* and property *G* are distinct if and only if they have different quiddities. *F*'s and *G*'s possession of different quiddities ensures that *F* and *G* are different properties, and conversely, if *F* and *G* have the same quiddity, then $F = G$.

Quidditism has some striking implications. Just as haecceitism implies that individuals have no essential properties other than their haecceities, quidditism implies that properties have no essential causal powers. Properties do not play their power-conferring roles essentially. Any power-conferring work they do is done in conjunction with something else, such as laws of nature. On Armstrong's view, for instance, if instances of *F* regularly produce instances of *G*, this is due to a law according to which *F*s necessitate *G*s.[1] This law is itself an instance of a property, in particular, a necessitation relation, *N*, that is instantiated in this world by the properties *F* and *G*. As a result, any instance of *F* in this world produces an instance of *G*. The power that instances of *F* have in this world to produce instances of *G* is thus accounted for by

[1] Fred Dretske (1977) and Michael Tooley (1977) have independently developed similar accounts of laws.

appeal to a law of nature. That law, importantly, must be metaphysically contingent, for if quidditism is true, there are no metaphysically necessary connections between properties and their powers—between, for instance, *F* and *G*, on the one hand, and the power of *F* to produce *G* on the other. *F* and *G* might instantiate *N* in the actual world, but they need not do so. In other possible worlds *F* might not necessitate *G* at all; it could be the case that *G* instead necessitates *F*, or that *G* and *F* are entirely unrelated. It is entirely possible, for instance, for round objects not to have the power to produce concave depressions in soft surfaces; roundness need not play that causal role.

Quidditism and the corresponding account of laws face several challenges. I'll describe some of them. First, if quidditism is true, then it is possible that a property could in principle take on any causal role whatsoever (Shoemaker 1980; Black 2000; Mumford 2004; Bird 2007). Shoemaker states the gist of the argument as follows:

> Suppose that the identity of properties consisted of something logically independent of their causal potentialities. Then it ought to be possible for there to be properties that have no potential whatever for contributing to causal powers . . . Further, it ought to be possible that there be two or more different properties that make, under all possible circumstances, exactly the same contribution to the causal powers of the things that have them. Further, it ought to be possible that the potential of a property for contributing to the production of causal powers might change over time . . . Thus a thing might undergo radical change with respect to its properties without undergoing any change in its causal powers, and a thing might undergo radical change in its causal powers without undergoing any change in the properties that underlie these powers. (1980: 214–15)

Bird (2007: 71–6) brings out the absurd implications of the contingency of causal roles by imagining a world, *w*, in which two properties, mass and negative charge, say, swap the causal roles they have in the actual world; that is, in *w* mass plays all the causal roles of negative charge, and negative charge plays all the causal roles of mass. Clearly, such a world would by all appearances be indistinguishable from the actual world. Intuitively, it seems odd to claim that the difference between *w* and the actual world is a difference in the causal roles of mass and negative charge. It seems much more plausible to claim that in world *w*, mass—the very property of mass that we are familiar with in the actual world—is called 'negative charge,' and negative charge—the very property of negative charge that we are familiar with in the actual world—is called 'mass.' What's been swapped in *w* are not the causal roles of mass and negative charge, but simply their names. The properties themselves, along with their causal roles, do not differ from the actual world.

The foregoing metaphysical worry is accompanied by an epistemological one. The metaphysical worry concerns the possibility that the same property might play different causal roles; the epistemological worry concerns the possibility that different properties might play the same causal role. Here is the way Shoemaker describes it:

> If there can be properties that have no potential for contributing to the causal powers of the things that have them, then nothing could be good evidence that the overall resemblance

between two things is greater than the overall resemblance between two other things; for even if A and B have closely resembling effects on our senses and our instruments while C and D do not, it might be (for all we know) that C and D share vastly more properties of the causally impotent kind than do A and B. Worse, if two properties can have exactly the same potential for contributing to causal powers, then it is impossible for us even to know . . . that two things resemble one another by sharing a single property. Moreover . . . it is impossible for us to know . . . that something has retained a property over time, or that something has undergone a change with respect to the properties that underlie its causal powers. (1980: 215)

To bring out this worry, imagine that in world w^* there are two different properties, inertial mass_A and inertial mass_B, which both play the causal role that defines inertial mass, the role expressed by '*m*' in the equation '*F=ma*.' Inertial mass_A and inertial mass_B play this role in different situations; in one situation a body accelerating in response to a force is due to inertial mass_A, while in another situation it is due to inertial mass_B. Since inertial mass_A and inertial mass_B play the same causal role, there is no way to tell in any particular case which is operative. If this seems an absurd result, Bird (2007: 77–9) argues that matters are even worse than they initially appear. The reason is that the definitions of theoretical terms are typically taken to specify unique referents (Lewis 1970). Inertial mass, for instance, might be defined as the (one and only) property that plays in a world the causal role expressed by '*m*' in the equation '*F=ma*,' or as the (one and only) property that plays in the actual world the causal role expressed by '*m*' in the equation '*F=ma*.' But a world like w^*, in which two properties play the causal role expressed by '*m*' in the equation '*F=ma*,' is a world in which nothing satisfies this definition, for it is a world in which there is not exactly one property that plays the causal role expressed by '*m*' in the equation '*F=ma*.' In w^*, therefore, the term 'inertial mass' fails to refer. Moreover, we do not know that the actual world is not like w^*; if quidditism is true, there is no way to tell. Consequently, if quidditism is true, we do not know whether our theoretical terms actually refer to anything at all. But surely this is a completely unacceptable conclusion. It entails among other things that it is impossible for us to know what properties really exist.

Quidditists have responded to these arguments by biting the bullet. With regard to the possibility that a property could play radically different causal roles from those it plays in fact, Armstrong says, "I think that the universals theory has to accept the case. Such a scenario is a possibility. But it is a *mere* possibility" (2010: 52). Lewis (2008), moreover, countenances the foregoing skeptical implication under the heading of Humility, the claim that we cannot know the fundamental properties of things. Lewis' argument for Humility is worth considering since it brings out a further argument against quidditism.

Lewis' argument for Humility is roughly as follows: Let T be a theory that introduces the terms $t_1, t_2, \ldots, t_n$ to refer to properties. T expresses the logical implications of $T(t_1, t_2, \ldots, t_n)$, a sentence Lewis calls the 'postulate of T.' If we replace the terms $t_1, t_2, \ldots, t_n$ in the postulate with variables $x_1, x_2, \ldots, x_n$, and quantify over the resulting expression, we get the Ramsey sentence of T: $(\exists x_1)(\exists x_2) \ldots (\exists x_n)T$

($x_1, x_2, \ldots, x_n$). Roughly, the Ramsey sentence expresses the theoretical roles played by the properties referred to by $t_1, t_2, \ldots, t_n$. Lewis argues that even though we can know about these roles, we cannot know which properties actually occupy them since it is possible for different properties to occupy a given role. This is an implication of quidditism. For any given role, x_i, specified by the Ramsey sentence of T, there are multiple properties which could occupy x_i, and no way of knowing which property occupies it in fact. Properties are thus analogous to Lockean substances on the quidditist view: they are unknowable we-know-not-whats that occupy theoretical roles. By contrast, dispositional essentialism posits no such inscrutable entities. It instead takes the theoretical roles of properties to constitute the essences of those properties. Seen from this perspective, the properties postulated by quidditists appear to be ontological extravagances, needless additions that are better done without.

The views of laws that accompany quidditism are also problematic. These include regularity theories of laws like Lewis' (1973a) and necessitation theories of laws like Armstrong's (1983). Regularity theories are based on Humean supervenience, the thesis that the world consists entirely of "local matters of particular fact, just one little thing and then another" (Lewis 1986b: ix). Laws of nature, on this view, are taken to be regularities among these matters of fact, in particular, regularities implied by every true deductive systematization of those facts that achieves an optimal balance of simplicity and strength (Lewis 1973a: 73). Insofar as local matters of fact consist of individuals having categorical properties, the laws of nature at a world supervene on the distribution of individuals and properties at that world. There are several well-rehearsed objections to this view of laws (Mumford 2004; Bird 2007). I won't describe them in detail. I'll simply mention that they include the problem of accommodating probabilistic laws, which Lewis calls the "one big, bad bug" of Humean supervenience (1986b: xiv). In addition, laws are supposed to explain the regularities we observe in the world, yet it is difficult to see how laws could explain regularities if they themselves are those regularities, for it seems implausible to suppose that regularities could explain themselves (Armstrong 1983: 40–1; Bird 2007: 86–90).

Necessitation theories like Armstrong's (1983) claim that laws are relations between universals; that is, they are higher-order universals, ones that are instantiated by other universals. If it is a law in world w that Fs produce Gs, there is a necessitation relation, N, which is contingently instantiated in w by the pair of universals F and G. The state of affairs $N(F,G)$, in other words, obtains in w, though it may not in other possible worlds. This view too faces several well-rehearsed objections. First, there is a basic problem characterizing the necessitation relation (van Fraassen 1989: chapter 5; Bird 2007: 91–8; Mumford 2004: 101–4). How exactly does the necessitation relation, N, which is a relation between the universals F and G, make a difference to how matters stand among particular instances of F and G? Necessitation theorists, it seems, must provide an account of what N is that explains how N is able to do the necessitating and explaining work that laws of nature are

supposed to do, yet no account seems to be forthcoming. Armstrong looks to deal with the problem by positing necessitation as a primitive:

> [T]he following complaint may be made. At the end of all our explanations, this factor of necessitation remains unexplained. This reproach is just, I think, but the inexplicability of necessitation just has to be accepted. Necessitation . . . is a primitive . . . which we are forced to postulate. (Armstrong 1983: 92)

Yet, it is not clear that this provides an adequate response to the problem. Consider the response Lewis makes in a well-known passage:

> What leads me . . . to reject Armstrong's theory . . . is that I find its necessary connections unintelligible. Whatever *N* may be, I cannot see how it could be absolutely impossible to have *N*(*F*,*G*) and *Fa* without *Ga*. (Unless *N* just is constant conjunction, or constant conjunction plus something else, in which case Armstrong's theory turns into a form of the regularity theory he rejects.) The mystery is somewhat hidden by Armstrong's terminology. He uses 'necessitates' as a name for the lawmaking universal *N*; and who would be surprised to hear that if *F* 'necessitates' *G* and *a* has *F*, then *a* must have *G*? But I say that *N* deserves the name of 'necessitation' only if, somehow, it really can enter into the requisite necessary connections. It can't enter into them just by bearing a name, any more than one can have mighty biceps just by being called 'Armstrong.' (1983b: 40)

In addition to the foregoing problem, which strikes at the core of Armstrong's account, there is another. According to Armstrong, there are no uninstantiated universals. If it is a law that spherical objects leave concave depressions in soft surfaces, then the necessitation relation that connects being spherical to leaving concave depressions must be instantiated by those universals, and those universals must be instantiated in turn; there must be at least one spherical individual that leaves a concave depression in something. Recall, however, that Armstrong looks to countenance talk of powers by appeal to laws of nature. To say that spherical individuals have the power to leave concave depressions is to say that there is a law of nature that connects sphericity to leaving concave depressions. Since this relation and the universals it connects must be instantiated, it follows that there can be no unmanifested powers on Armstrong's view; there must be at least one case in which a power is manifested. Yet, this is a highly counterintuitive result. Consider again Martin's (1996a, b, c) example: There are two kinds of subatomic particles, A-particles and B-particles, which have properties conferring on them the power to interact with each other. Nevertheless, no A-particles and B-particles exist near enough one another to ever actually interact. Intuitively, this kind of situation seems possible, yet on Armstrong's view it is not, for A-particles and B-particles could not have the power to interact unless some A-particle actually interacted with some B-particle.[2]

[2] Heil (2003: 90–2) describes a related problem: Armstrong's view implies that the counterfactual conditionals supported by disposition ascriptions are false when they should be true.

Before moving on to other competing accounts of powers and properties, it's worth noting that the conjunction of quidditism with trope theory is particularly problematic. Tropes must be able to resemble each other if they are to account for objective similarities among individuals. But in respect of what can they resemble each other if not their theoretical roles, the very theoretical roles that according to quidditism properties do not have essentially? To appreciate the difficulty here, suppose that quidditism is true and that properties are tropes. Suppose, moreover, that the resemblance class of *F*-tropes comprises only the tropes F_1 and F_2. Since quiddities are primitive principles of identity and distinctness, they don't appear to play any role other than ensuring that each property is identical to itself and distinct from everything else. But this theoretical role seems incapable of grounding the resemblance class of *F*-tropes. The quiddity of every property must serve to distinguish it from every other property, but in that case, the quiddity of every trope will be different from the quiddity of every other. F_1's quiddity will not resemble F_2's, for F_1's quiddity must ensure that F_1 is not F_2, whereas F_2's quiddity cannot ensure that F_2 is not F_2. Quidditists might argue in response that all quiddities resemble each other in their basic form: the quiddity of any property, *P*, ensures that *P* is itself and nothing else, and in that sense, the quiddities of F_1 and F_2 can ground a resemblance between F_1 and F_2. The problem with this proposal is that the resemblance is too broad, for not just F_1 and F_2 will resemble each other in this respect, all properties will. Since the quiddities of all properties will have this general form, the corresponding resemblance class will include all tropes, not just *F*-tropes. Since there is no more to a quiddity than being a principle of identity and distinctness, and since quidditism implies that there is no more to the nature of a property than its quiddity, it appears that quiddities cannot ground any but the very broadest resemblances among properties. But if those properties are tropes, we are left with a theory of properties that cannot account for most of the relevant objective similarities among things, and that is not a workable theory of properties at all.

The foregoing arguments show that attempts to understand powers in terms of categorical (that is, nondispositional) properties together with laws of nature are problematic. They are not knockdown arguments for a variety of reasons (Mumford 2004), but exponents of dispositional essences can nevertheless argue that their view is more plausible on balance than the categoricalist alternative.

5.3 Higher-Order Properties

Another view that posits categorical properties claims that dispositions are higher-order properties that have lower-order categorical properties as their bases (Prior et al. 1982). Calling a property 'higher-order' is not the same as calling it 'higher-level.' Levels can be introduced using any number of principles for hierarchical organization, but to speak of higher- and lower-orders is to speak of a specific kind of hierarchical principle. There are, moreover, two different uses of the term 'higher-order.' According

to one of them, a higher-order property is a property had by another property. It is in this sense of the term that Armstrong claims that necessitation is a higher-order relation; it is a universal instantiated by other universals.[3] But this is not the sense that concerns us here. Higher-order properties in the sense that concerns us here are supposed to be properties of individuals. Higher-order properties in this sense are logical constructions that are defined using the logical operation of quantification. They are properties whose definitions quantify over other properties. If $F_1, F_2, \ldots, F_n$ are properties, we might define a property, H, by stipulating that something has H if and only if it has *some* F-property or other. This definition quantifies over F-properties; H is thus a higher-order property. It is not the property of a property; it is rather the property of an individual having a property.

The definitions of higher-order properties are usually not as simple as the example of H suggests since those definitions usually specify a condition on the lower-order properties over which they quantify. We might define H not simply as the property of having some F-property, but as the property of having some F-property that satisfies condition C. Hilary Putnam (1970: 313–14) once suggested that mental properties were higher-order properties of this condition-specifying sort. Having a mental property amounted to having some first-order property that was related to other first-order properties in ways that satisfied a certain functional description. Being in pain, for instance, might be defined as having some first-order physical property that correlates pinpricks, burns, and abrasions with winces and groans, and various internal states. The idea that mental properties are higher-order properties is the basis of realization physicalism, a species of nonreductive physicalism (Jaworski 2011: 144–9). Prior et al. (1982) make an analogous claim about dispositions. According to them, a dispositional property such as fragility is the property of having some property (the causal basis of the disposition) that is responsible for, say, making its possessor shatter when dropped. Theirs is a categoricalist view like Armstrong's and Lewis' insofar as they take the properties over which the definitions of dispositional properties quantify to be categorical. Notice, moreover, that the higher-order properties postulated by realization physicalists and by Prior et al. are supposed to be properties of individuals, not properties of properties. Fragility is supposed to be a property of the individual vase, and being in pain is supposed to be a property of an individual person.

There are at least two problems facing a higher-order account of powers like Prior et al.'s. First, because it is committed to there being purely categorical properties, it would appear to face the problems that any categoricalist theory faces. We saw in the

[3] This is the sense of 'higher-order' that Molnar (2003: 30–1) has in mind when he argues that there are higher-order properties. He defends the claim that there are properties had by other properties; he does not defend the claim that there are properties which are mere logical constructions. As I argue below, such properties could make no causal difference to their bearers, and consequently claiming that they exist would violate the Eleatic principle, which Molnar himself accepts.

previous section that these include problems with quidditism and also problems with the corresponding theories of laws, whether they be regularity theories or necessitarian ones. Second, the higher-order account of dispositions faces a problem accounting for the causal efficacy of powers. A line of reasoning made familiar in the philosophy of mind by Jaegwon Kim (1992, 1993b, 1998) concludes that higher-order properties are epiphenomenal: they do not confer causal powers on their possessors. The Eleatic principle introduced in Section 2.2, however, implies that the only properties that exist are ones that empower individuals to enter into causal relations. If higher-order properties are epiphenomenal, if they do not empower their possessors, then the Eleatic principle implies that higher-order properties do not exist. Let us consider this argument in detail.

Let *H* be a higher-order property. It could be a mental property as realization physicalists claim, or a dispositional property as Prior et al. claim. Suppose that *H* is realized at time *t* by a first-order property, *F*. This is typically taken to imply that *H* is not identical to *F* since it is possible for *H* to be realized by other first-order properties. Realization physicalists argue for this point by appeal to the multiple-realizability of mental properties, and we saw in Section 4.5 that Prior et al. (1982: 253–4) argue for a similar thesis about dispositions: fragility could be correlated with molecular bonding α in one case, and with crystalline structure β in another; consequently, they say, fragility can be identical to neither α nor β since identity implies necessary coextension.

Kim argues that if *H* is realized by *F* at time *t*, then *H* can have no causal powers at *t* other than those of *F*. He sometimes calls this premise 'Causal Inheritance' (Kim 1992: 326; 1998: 54):

> If a higher-order property *H* has a first-order property *F* as its realizer (what Prior, Pargetter, and Jackson call its 'causal basis') at time *t*, then this instance of *H* cannot have any causal powers at *t* other than the powers of this instance of *F* at *t*.

Kim argues that denying Causal Inheritance is tantamount to believing in magic:

> For there is nothing in the instantiation of [the higher-order property] . . . over and above the instantiation of its realizer . . . [T]o think that [the higher-order property] has causal powers in excess of those of [its realizer] is tantamount to belief in magic: somehow new causal powers miraculously materialize as we define second-order and higher-order properties from a fixed stock of first-order properties! (1998: 54–5)

To appreciate Kim's point, suppose that $F_1, F_2, \ldots, F_n$ are all the first-order properties there are. Let us now introduce a new way of talking about these properties; let us define the predicate '*H*' by stipulating that something has *H* if and only if it has some *F*-property or other that satisfies condition *C*. By doing this, we do not add to the basic inventory of properties that exist. We cannot bring about the existence of new first-order properties by sheer fiat, the way we can imagine God, say, bringing about the existence of a new fundamental physical property to add to $F_1, F_2, \ldots, F_n$.

Defining a predicate like 'H' does not add to the basic features of the world, it merely introduces a new way of talking about them. By analogy, deciding to call players on the team who satisfy one condition 'infielders' and those who satisfy another condition 'outfielders' does not add players to the roster. Suppose, then, that a has F_1 at t, and that a's having F_1 at t satisfies condition C. In that case, the definition of H implies that a has H at t as well. But simply applying the predicate 'H' to a does not magically bring into existence any new causal powers. Whatever causal powers a has are conferred on it by the first-order properties it has, they do not depend on the applicability of the predicate 'H.' There is thus good reason to think that Causal Inheritance is true. If H is a higher-order property that is realized by a first-order property F at time t, then this instance of H cannot have any causal powers other than those of F.

It's noteworthy that nonreductive physicalists have sometimes touted the implications of Causal Inheritance as an asset of their view. The reason, they say, is that it enables them to endorse antireductionism without compromising their commitment to physicalism. Suppose that $F_1, F_2, \ldots, F_n$ are all the properties postulated by fundamental physics. We can nevertheless use higher-order predicates, such as mental predicates, to satisfy descriptive and explanatory interests that the predicates of fundamental physics do not enable us to satisfy. Because the latter predicates do not enable us to satisfy the interests that the former predicates do, fundamental physics cannot take over the descriptive and explanatory roles that higher-order discourse plays. But if fundamental physics cannot take over the descriptive and explanatory roles of higher-order discourse, then higher-order discourse is not reducible to fundamental physics since reduction consists precisely in one theoretical framework taking over the descriptive and explanatory roles of another. The view is thus committed to antireductionism. Yet, because higher-order predicates do not add to the basic inventory of properties that exist, nonreductive physicalists are still free to claim that the only properties that exist are those postulated by fundamental physics. They thus conclude that antireductionism is compatible with physicalism.

If Causal Inheritance is true, then higher-order properties do not confer any distinctive causal powers on the things having them. Again, by analogy, if the number of players on the roster increases, this cannot be due simply to introducing the terms 'infielder' and 'outfielder.' Consequently, if mental properties and other properties postulated by the special sciences are higher-order, they cannot add to the causal powers that already exist. If Prior et al. are right, moreover, then the same is true of dispositions. The causal basis of a dispositional property, they say, is sufficient to account for any effects attributed to the disposition, but there cannot be multiple sufficient causes; consequently, the higher-order dispositional property can make no causal contribution to the effect. The vase's shattering is not due to its being fragile; it is due instead to the causal basis of fragility, a lower-order property such as having molecular bonding α. This and this alone explains why the vase shatters when dropped.

According to the Eleatic principle introduced in Section 2.2, however, the only things that exist are ones that play some sort of causal role. This implies that the only properties that exist are ones that empower individuals to enter into causal relations. But the mental properties of realization physicalists and the dispositions of Prior et al. do not enable individuals to enter into causal relations. The Eleatic principle thus implies that these so-called properties do not exist. What exist instead are predicates: mental predicates such as 'is in pain' and dispositional predicates such as 'is fragile.' These predicates may express properties, but the properties they express are their first-order realizers or causal bases. If the predicate 'is fragile' applies to this vase on account of the vase's having molecular bonding α, then in this particular case, the predicate 'is fragile' expresses that property. There is no further, higher-order property, *being fragile*, which is distinct from this first-order realizer.

Why do some philosophers speak as if there is such a property? The reason is that they claim that properties are abundant, not sparse. According to Prior et al., for instance, "*every* expression like 'the property of being *F*' denotes a property" (1982: 254, emphasis in original). Consequently, the very fact that we use the predicate 'is fragile' or the term 'the property of being fragile' commits them to fragility being a property. We have already seen, however, that if properties are sparse, this conclusion does not follow. If we remain committed to sparse properties, then we must forgo talk of higher-order properties. Rather, as Kim (1998: 104) suggests, it is much clearer to speak of higher-order predicates, concepts, or descriptions than it is to speak of higher-order properties (cf. Heil 2003: 45). The latter term is misleading. There are no higher-order properties if properties are sparse, and as we saw in Chapter 2, there are good reasons to think that properties are sparse, especially when dealing with issues in the philosophy of mind. In the meantime, let us return to the higher-order account of dispositional properties endorsed by Prior et al.

Based on what's been said, Prior et al.'s view can be characterized as the view that there are no dispositional properties, but only dispositional predicates that we use to express categorical properties such as the vase's having molecular bonding α. It is categorical properties like this that are responsible for the changes things undergo, such as the vase's shattering when dropped. But how exactly do categorical properties explain these changes on this view? Prior et al. cannot claim that the categorical properties are identical to powers, for in that case, the so-called categorical properties would not be categorical, they would be dispositional. Yet, their view shuns dispositional properties. It seems that the only alternative is to endorse one of the views already considered, either a view that offers conditional analyses of disposition ascriptions, or a view like Armstrong's or Lewis' which conjoins categorical properties with laws. We've seen, however, that there are good reasons to find these views unsatisfactory. Consequently, it seems that we have to draw a similar conclusion about the higher-order property view.

5.4 Dualistic Theories

Finally, some views conjoin a commitment to categorical properties with a commitment to pure dispositions. Ellis (2001, 2002) and Molnar (2003) both endorse views of this sort. Since there are problems associated both with categorical properties and with pure dispositions, these views would appear to inherit the difficulties associated with both. If those difficulties give us good reason to reject categorical properties and pure dispositions, as I've argued, then they also give us good reason to reject views that are committed to both. In addition, there is some reason to suspect that there are further problems with views committed to both when it comes to explaining how categorical and dispositional properties are related (Armstrong 2002: 169–71; 2005: 313).

5.5 Conclusion

In Chapter 4, I articulated and defended the identity theory of powers. In this chapter, I've criticized competing theories of powers, categoricalist theories chief among them. We saw that these theories are committed to quidditism, the claim that properties have brute principles of identity and individuation, and that quidditism is typically combined with an account of metaphysically contingent laws of nature. We saw that both of these commitments encounter difficulties. In addition, categoricalists like Prior et al. take dispositions to be higher-order properties, properties whose definitions quantify over other properties. They argue that these higher-order properties make no causal contribution to things beyond the contributions made by their lower-order categorical bases. If they are right, then there is good reason to think that higher-order properties do not exist, for the Eleatic principle introduced in Section 2.2 implies that the only existing properties are ones that enable individuals to enter into causal relations. But the arguments that Prior et al. advance, and an argument to the same effect that Jaegwon Kim advances, imply that higher-order properties to do not enable individuals to enter into causal relations. Any causal work that might be attributed to so-called higher-order properties is done instead by their lower-order causal bases (what Kim refers to as their lower-order realizers). The claim that there are no higher-order properties will become important later on when discussing nonreductive physicalism (Section 11.4).

The goal of the last few chapters has been to articulate and defend a metaphysical framework that will enable us to discuss hylomorphism and its competitors with minimal terminological confusion. Now that that framework is in place, we are in a position to situate the hylomorphic notion of structure within it.

6

Structured Individuals and Their Parts

6.1 What Are Structures?

The goal of Chapters 2–5 was to develop a metaphysics that we could use to frame a discussion of hylomorphism and its implications for the philosophy of mind. That metaphysics takes individuals to be the primary agents, and it takes their properties to be powers to act and be acted upon. The goal of this chapter and the following two chapters is to situate an account of structure within this metaphysics.

According to hylomorphists, some individuals, paradigmatically living things, are composed of physical materials with structures. We can call the structures that make individuals what they are *individual-making structures*. Individual-making structures are not the only ones that exist, according to hylomorphists. There are also *activity-making structures*, structures that make the activities of structured individuals what they are. Activity-making structures are the subject of Chapter 8. This chapter and the next, however, focus on individual-making structures, and when I speak of structures in these chapters, it should be taken to refer to individual-making structures, unless otherwise noted. The goal of this chapter is to describe the notion of individual-making structure and its relation to the hylomorphic account of composition. The goal of Chapter 7 is to defend that account against some objections.

Consider again the sketch of the hylomorphic worldview presented in Chapter 1. As we gaze out at the vast sea of matter and energy that is or will be described by our best physics, we see in it numerous localized pockets of organization or order—semi-stable, self-maintaining warps or vortices of physical material. According to hylomorphists, each of these vortices is a distinctive individual, paradigmatically an organism. To understand one of them, we need to understand at least two things about it. First, we need to understand the physical materials that are caught up into it. Those materials, their characteristics, and behavior, will tell us something about the characteristics and behavior of the structured individual as a whole. But those materials will not by themselves provide an exhaustive understanding of it, say hylomorphists. They will not, for instance, enable us to understand what sets the structured whole apart as a unified individual distinct from the inanimate materials that surround it. Physical materials, including those that are currently caught up into

the structured individual, can exist without being caught up into it or any other structured whole, and any given structured whole can exist without the materials that are caught up into it at any given moment. Consequently, say hylomorphists, there must be some further principle that accounts for the unity and also the persistence of structured wholes, a principle that explains why numerically one and the same whole can exist at different times despite a constant influx and efflux of materials that compose it. In addition, say hylomorphists, that principle should explain why the whole has powers not had by the materials that compose it, such as the powers to think, perceive, and feel. According to hylomorphists, understanding the unity, persistence, and distinctive powers of these wholes requires understanding their structures.

What exactly are structures? The metaphysic developed in Chapters 2–5 implies that they must be either individuals or properties, and the theoretical roles we expect structures to play are characteristic of properties. Structures, for instance, are supposed to confer powers. The squashing example introduced in Section 1.1 suggests that Godehard's structure is what confers on him the powers to think, feel, perceive, and act. According to the metaphysic defended in Chapters 2–5, it is properties that confer powers. Consequently, if we assume that metaphysic, we appear committed to structures being properties.

If structures are properties, then they have all the characteristics of properties discussed in Chapters 3 and 4. First, they must be powers—powers in particular to configure (or organize, order, or arrange) materials. Each structured individual organizes or configures the materials that compose it. I configure the materials that compose me, and you configure the materials that compose you. Describing the way each of us configures our respective materials is something that hylomorphists say is an empirical undertaking—in our cases, an undertaking left largely to biology, biochemistry, neuroscience, and other biological subdisciplines. Collectively, these disciplines are likely to deliver long, complicated descriptions of cells, tissues, and organ systems, along with their characteristic activities, capacities, and interrelations. It will be convenient to have a term to stand in for these descriptions. Let us say that you and I configure materials *human-wise*, where 'human-wise' is a placeholder for the longer descriptions that it is the collective job of biologists, neuroscientists, and others to provide.

Second, structures are particulars. To say that you and I configure materials human-wise does not imply that there is a universal, *configures human-wise*, that you and I have in common. The trope theory developed in Chapter 3 implies instead that my configuring and your configuring are numerically different properties. My configuring and yours nevertheless resemble each other rather closely—more closely than, say, either resembles Fido's configuring the materials that compose him or the oak tree's configuring the materials that compose it. The trope theory also implies that my configuring and yours are nontransferable in the sense described in Section 3.2: my configuring cannot belong to anything other than me, nor can your configuring belong to anything other than you.

Third, structures have the same directedness that all powers do. The structures of living things in particular appear to be directed toward developing and maintaining organisms' mature states, as well as the powers that characterize those states and their manifestations. Ernst Mayr describes the idea this way:

Living organisms are complex, ordered systems. This explains their capacity for regulation and for control of the interaction of the genotype, as well as their developmental and evolutionary constraints ... [They] are programmed for teleonomic (goal-directed) activities from embryonic development to the physiological and behavioral activities of the adults ... [They] go through a definite life cycle beginning with a zygote (fertilized egg) and passing through various embryonic or larval stages until adulthood is reached. (1997: 22)

Like many of the powers we have considered, an organism's structure manifests itself in many different ways—in, for instance, the various developmental stages Mayr mentions, as well as in the variety of self-regulating processes that maintain the cells, tissues, and organs living things develop: photosynthesis, glycolysis, protein synthesis, and so on.

Mayr's comments bring out an important point first introduced in Chapter 1: some structures are *dynamic*. They are not relatively stable spatial relations among an individual's parts; they instead have a temporal dimension: they comprise sequences of changes over time. They often have a conditional dimension as well: they involve different kinds of changes under different kinds of conditions. Moreover, individuals with dynamic structures typically survive changes in the materials they structure. Recall Jonathan Miller's remarks, which were quoted in Chapter 1:

[T]he survival of form depends on ... [either] the intrinsic stability of the materials from which the object is made, or the energetic replenishment and reorganisation of the material which is constantly flowing through it ... [A] fountain ... can retain its shape only by endlessly renewing the material which constitutes it; that is, by organising and imposing structure on the unremitting flow of its own substance ... The persistence of a living organism is an achievement of the same order as that of a fountain ... it can maintain its configuration only by flowing through a system which is capable of reorganising and renewing the configuration from one moment to the next. (1978: 140–1)

Kit Fine and Mark Johnston, two fellow hylomorphists, also accommodate the idea of dynamic structures. Fine (1999: 68–9) calls them 'principles of variable embodiment'; Johnston calls them 'dynamic principles of unity':[1]

A form or principle of unity may be ... *dynamic*, in that ... the parts it holds of vary over time ... A paradigm case is a living thing whose organic matter is unified into an organism by some categorical basis of a multi-track disposition to such life-functions as ingestion, assimilation,

[1] Robert Koons, another hylomorphist, uses the term 'formal processes': "*Formal processes* ... have fundamental properties that are temporally extended—irreducible properties of motion and change. The instantaneous properties of the process and its participants are grounded in the temporally extended properties of the process" (2014: 166).

excretion, growth, metabolic repair, and so on and so forth. In this case, the principle of unity is a complex structure of biochemical relations, whose holding of the organic matter of the living thing provides the categorical basis for the multi-track disposition in question. The operation of that disposition *requires* the matter to be exchanged over time. (2006: 663–4)[2]

Some contemporary hylomorphists use the term 'structure' in a much narrower, static sense. We saw in Chapter 1 that David Oderberg is among them. He understands structure as a "configuration of parts" which, he says, is "far too *static* a concept to tell you all there is about the form of an animal" (2014: 177). This is an important terminological difference. What Oderberg calls 'forms' play the same kinds of theoretical roles that individual-making structures do; in particular, *forms count* on Oderberg's view: they operate as principles of unity, just as individual-making structures do on the hylomorphic theory I've been developing. Unlike Oderberg, I intend to use expressions such as 'spatial arrangements' or 'mere spatial arrangements' to refer to spatial relations among something's parts. The term 'structure' I will reserve for properties that play the theoretical roles mentioned in Chapter 1: *structures matter*, *structures make a difference*, and *structures count*.

A fourth feature of structures is that they confer whatever powers they do necessarily. It is metaphysically impossible for my structure not to confer on me the power to grow lungs, skin, and bones. Critics, especially those sympathetic to the idea that laws of nature are contingent, might object that we can easily conceive of situations in which I fail to develop normally, and end up without lungs, or skin, or bones. We saw in Section 4.3, however, that there are at least two ways hylomorphists can respond to this kind of objection. Either they can deny that worlds in which I lack the power to develop lungs, skin, and bones are genuinely conceivable, or else they can deny that the kind of conceivability we achieve in these cases is a reliable guide to possibility. Consider an example of the first strategy. Hylomorphists can claim that when we take ourselves to be conceiving of me lacking the power to develop lungs, skin, or bones, we are really conceiving of me in circumstances in which something is inhibiting my power to develop normally. Powers, recall, can inhibit or enhance the manifestation of other powers (Section 4.2)—that includes the powers of living things. It is possible for outside influences to interfere with the ways my structure is manifested—teratogenic agents that could prevent me from developing lungs, skin, or bones. If critics insist that they are conceiving of me failing to develop lungs, skin, or bones in the absence of outside influences, hylomorphists have the option of shifting to the second strategy: they can accept that the critics are in some sense conceiving what they say, but deny that conceiving in this sense is any guide to what

[2] The account of structures I defend nevertheless differs in significant ways from Fine's and Johnston's. One difference that comes out in this passage is that Fine and Johnston conceive of hylomorphic structures as relations among something's parts. I conceive of them rather as relations between wholes and their parts: a whole configures or structures its parts. This has important implications for how my hylomorphic theory avoids some versions of Williams' worry (section 12.2). I discuss other differences between the hylomorphic theory I defend and the theories defended by Fine and Johnston in Chapter 14.

is possible. Perhaps critics are tacitly employing an epistemic notion of conceivability. When they say that they are conceiving of me failing to meet various developmental milestones in the absence of any noxious outside influences, they are merely saying that such a circumstance is not inconsistent with everything they know about human development. This kind of conceivability is clearly not necessarily an indicator of possibility, for among other things, critics might not know very much about human development. It is possible that if they knew more—if, for instance, they knew about the causal basis of human development (Section 4.4)—they might no longer be able to conceive what they currently do.

With these points in mind, consider again the theoretical roles attributed to structures in Chapter 1. We expressed them in four slogans:

Structure matters: it operates as an irreducible ontological principle, one that accounts at least in part for what things essentially are.

Structure makes a difference: it operates as an irreducible explanatory principle, one that accounts at least in part for what things can do, the powers they have.

Structure counts: it explains the unity of composite things, including the persistence of one and the same living individual through the dynamic influx and efflux of matter and energy that characterize many of its interactions with the wider world.

Structure minds: it provides us with resources for understanding the place of mental phenomena within the natural world.

We are in a position to consider the first two of these roles; the third we'll consider in Sections 6.2 and 6.3, and the fourth in Chapters 8–13.

Based on what's been said, it should be apparent how individual-making structures make a difference on the hylomorphic view. Structures are powers. Consequently, if something has an individual-making structure, it has powers that it would not otherwise have. Individual-making structure thus makes a difference. Individual-making structure also matters on the hylomorphic view. According to it, structured individuals have their configuring properties essentially: each is essentially an organizer/configurer of materials. For a structured individual to cease configuring some materials or other is for that individual to cease to exist. To use Furth's analogy (Section 1.3), a knot of matter and energy that unravels is no longer a knot. Individual-making structures are thus essential properties of structured individuals.

In addition, structures cannot go unmanifested; in this sense, they are unlike the other powers we have considered. Crystals of table salt can sit idly; their power to dissolve in water can remain forever unmanifested. But there is no sitting idle when it comes to my power to configure the materials that compose me. If I am not manifesting that power, if I am not configuring those materials, then I do not exist, and if I do not exist, then there is no individual to do my configuring.

One helpful way of understanding the essential manifestation of structures is to think of them as reflexive powers. A structure is its own reciprocal disposition partner: it is its own condition for manifestation. Consequently, it manifests itself

at every time at which it exists.[3] Individual-making structures are not just essential powers of structured individuals; they are essential powers of structured individuals that are essentially manifested; they cannot exist unmanifested.

Structure is also supposed to count on the hylomorphic view; it is supposed to explain the unity and persistence of composite individuals, and in the case of living things, that means explaining their unity and persistence through the dynamic influx and efflux of matter and energy that characterize their interactions with the surrounding world. To understand how individual-making structure plays this role, we need to understand the hylomorphic account of composition. Describing and defending it will occupy the remainder of this chapter and the next.

6.2 Van Inwagen on Composition and Lives

A robust literature has grown up around the topic of composition due in large measure to the work of Peter van Inwagen (1990). Among familiar views of composition, the hylomorphic view resembles most closely his own. In what follows, I'll present the hylomorphic view as a way of developing van Inwagen's. I'll begin with a brief overview of the latter.

Van Inwagen presents his account of composition as an answer to the Special Composition Question (SCQ): Under what conditions do multiple things compose one thing? Or under what conditions is it true that there is a y such that the xs compose y?[4] Extreme answers to the SCQ are offered by mereological universalists and mereological nihilists. Universalists endorse unrestricted composition; they claim that composition occurs under any conditions: any things whatsoever compose one thing. They thus include in their ontologies commonsense entities such as humans, trees, tables, and mountains, but also many entities that defy commonsense, such as the object composed of the northern half of Mount Rainier and your big toe. Nihilists, on the other hand, claim that composition occurs under no conditions: no things whatsoever compose anything. Their ontologies admit no entities like the object composed of the northern half of Mount Rainier and your big toe, but neither do they admit commonsense entities such as humans, trees, tables, and mountains. The only things that exist according to nihilists are mereological atoms or simples, individuals with no proper parts.

Moderate answers to the SCQ endorse restricted composition: composition does occur (contrary to nihilism), but it occurs only under certain conditions (contrary to universalism). Van Inwagen's own moderate answer to the SCQ claims that composition occurs exactly if the activities of some fundamental physical particles constitute a life. By 'a life' van Inwagen takes himself to mean what biologists do: "the individual life of a concrete biological organism ... [the] sense according to

[3] I am grateful to Joseph Vukov for bringing this idea to my attention.
[4] Here 'the *x*s' is a plural referring expression similar to 'the faculty members' or 'the infielders.'

which 'Russell's life' denotes a purely biological event" (van Inwagen 1990: 83). What exactly is such an event? According to van Inwagen, lives are self-maintaining events like flames and waves, except that, unlike flames and waves, they are *well-individuated* and *jealous*.

Flames are not as well-individuated as lives, van Inwagen argues: "If I light seven candles from one taper, has a spatially connected flame become a scattered flame, or have seven new flames come into existence? Presumably, there are no answers to these questions" (1990: 88). Waves are better individuated than flames, but waves for their part are not jealous:

> Consider two waves . . . which are moving in opposite directions and which pass through each other . . . I think we must say . . . that both the waves exist at the moment of superposition and that each is at that moment constituted by the activities of the same water molecules. We may describe . . . the possibility of two waves' being simultaneously constituted by the activities of the same objects . . . by saying that a wave is not a *jealous* event. Lives, however, are jealous. It cannot be that the activities of the *x*s constitute at one and the same time two lives . . . [I]t is not surprising that lives are jealous events. When two waves impinge upon the same water molecules, the activities that each demands of these molecules . . . sum neatly according to the rules of vector addition . . . A wave contributes energy to the particles of a fluid and then collects that same energy once more as it passes . . . A life, on the other hand, does not deposit and withdraw sequentially an invariant sum of energy . . . A life takes the energy it finds and turns it to its own purposes. (1990: 88–9)

Since lives are jealous, they cannot overlap the way waves do; there could not be a situation in which the activity of the *x*s constitutes a life, and the activity of the *y*s constitutes a life, and the *x*s are not identical with the *y*s, and for some *z*s, the *z*s are among both the *x*s and the *y*s. There is only one exception to this, says van Inwagen: it is the case in which one of the lives is *subordinate* to another, and the only cases of this sort, he thinks, are cases in which the life of an individual cell is subordinate to the life of a multicellular organism (1990: 89).

Beyond the foregoing points, van Inwagen's descriptions of lives stay largely at the level of metaphor and analogy. The reason is that providing the literal details about what lives are and what characteristics they have is, he thinks, a job for biologists. It is their business to answer the question 'What is life?', he says, "just as it is the business of chemistry to answer the questions 'What is a metal?' and 'What is an acid?', or the business of physics to answer the question 'What is matter?'" (1990: 84).

Van Inwagen is thus committed to a distinctive understanding of the tasks and limits of mereology—a distinctive meta-mereology. This understanding is largely implicit in what he says. Perhaps the closest he comes to stating it explicitly is the following passage:

> I have presented a certain picture . . . of the nature of a living organism . . . but it is not a philosopher's picture. That is, it is not a picture that could only come to be as a product of the attempts of philosophers to deal with their peculiar preoccupations. (1990: 92)

The reason van Inwagen's picture is not one that could result merely from philosophers' attempts to deal with their own preoccupations is that it relies on input from empirical sources. I take Kathryn Koslicki to give a fuller explication of the kind of meta-mereology van Inwagen gestures toward:

> I take the mereologist's job to be to devise an appropriate conception of parthood and composition which accurately reflects the conditions of existence, spatio-temporal location and part/whole structure of those objects to which we take ourselves to be already committed as part of the presupposed scientifically informed, commonsense ontology. The question of which kinds [of objects] there are I take to be . . . answered [not] by the mereologist proper, but by the ontologist at large, in conjunction with . . . science and common sense, which . . . have something to contribute to the question, "What is there?" . . . [M]ereology . . . does not settle matters of ontological commitment; rather, it presupposes them to be resolved elsewhere within metaphysics or outside of philosophy altogether . . . [This] approach differs from the standard conception [which is committed to unrestricted composition] as well as from Fine's theory of embodiments . . . which view the mereologist as a specialized sort of ontologist, whose job . . . is precisely to tell us what mereologically complex objects (if any) the world contains . . . [B]y presupposing that the question, "What mereologically complex objects (if any) are there?" is *descriptively* settled in the course of arriving at a scientifically and commonsensically acceptable ontology of kinds, the present approach assigns to the mereologist proper a more limited set of responsibilities . . . characteriz[ing] . . . those mereologically complex entities whose existence is already confirmed by independent evidence. (2008: 171)

Exponents of this kind of meta-mereology rely on empirical disciplines to provide starting points for philosophical reflection on what exists, including what kinds of parts exist. This point will become important later on when we consider hylomorphists' reasons for endorsing organic parts such as eyes, hearts, and hands.

Van Inwagen proposes that composition happens exactly when the activity of fundamental physical particles constitutes a life. His answer to the SCQ, in other words, is this: there is a y such that the xs compose y if and only if the activity of the xs constitutes a life. From this, he says, it follows that x is a proper part of y if and only if y is an organism and x is caught up in the life of y (1990: 94). The expression "caught up in a life" is one that van Inwagen borrows from the biologist J. Z. Young (1971). Van Inwagen explains with an example:

> Alice drinks a cup of tea in which a lump of sugar has been dissolved. A certain carbon atom . . . is carried along with the rest of the sugar by Alice's digestive system to the intestine. It passes through the intestinal wall and into the bloodstream, whence it is carried to the biceps muscle of Alice's left arm. There it is oxidized in several indirect stages (yielding in the process energy . . . for muscular contraction) and is finally carried by Alice's circulatory system to her lungs and there breathed out as a part of a carbon dioxide molecule . . . Here we have a case in which a thing, the carbon atom, was . . . caught up in the life of an organism, Alice. It is . . . a case in which a thing became however briefly, a *part* of a larger thing when it was a part of nothing before or after. (1990: 94–5)

So, according to van Inwagen, composition does not happen apart from lives; composite beings are all living things. Van Inwagen's view implies, then, that there are two kinds of material beings: mereological simples (material beings without proper parts) and living things, and the only living things, according to van Inwagen, are single cells and multicellular organisms.

Van Inwagen's view has several noteworthy implications. First, it implies what he calls 'the Denial,' the claim that many objects belonging to a commonsense ontology do not exist. Examples include artifacts such as tables and chairs, and also natural bodies such as rocks, mountains, stars, and planets. Since the only material beings on van Inwagen's view are living things and mereological simples, and since tables, chairs, mountains, and so on are neither simples nor living things, van Inwagen's view implies that these things do not exist (he calls them 'virtual objects'). Although this seems counterintuitive, van Inwagen argues that his view can accommodate our pedestrian intuitions. The reason is that it is possible to paraphrase statements that are apparently about artifacts and natural bodies in ways that mention only mereological simples. When we say that there is a chair in the corner, for instance, we are really saying that certain mereological simples are spatially arranged in the corner chair-wise. This strategy allows van Inwagen to countenance ordinary talk of artifacts, mountains, and stars without compromising his answer to the SCQ.

The Denial enables van Inwagen to solve a variety of philosophical problems in an elegant way. One is the Ship of Theseus. Plutarch reports that for many years the Athenians preserved the ship that Theseus had sailed to Crete by replacing the older planks with newer ones. It apparently became a source of debate among philosophers whether the ship with the replaced planks was in fact numerically identical to the original ship of Theseus. Thomas Hobbes later added a twist: suppose that the original planks, once discarded, were reassembled just as they had been in the original. Which ship, if either, would be the real ship of Theseus: the ship in the harbor composed of the replaced planks, or the ship assembled from the discarded ones? Van Inwagen's solution is to deny that there are any ships. If there are no ships, then there can be no question which ship is the real ship of Theseus.

Consider likewise the puzzle of the statue and the lump. The statue and the lump share all the same parts at exactly the same time. Surely two things cannot share all the same parts at the exactly the same time, so the statue and the lump must be identical. Yet, the statue and the lump cannot be identical since they appear to have different properties. The lump, for instance, existed before the statue, and unlike the statue, it can survive being squashed. In that case, though, it follows from Leibniz's law that the statue and the lump must be distinct. There are thus good reasons to think both that the statue and the lump are identical and that they are distinct. Van Inwagen's solution is to deny both that there are statues and that there are lumps. The particles that occupy the location that the statue or the lump would occupy if either existed compose nothing at all, neither a statue nor a lump. And if there is no statue and no lump, then there can be no question whether the statue and the lump are identical.

Part of what makes the foregoing puzzles interesting is that they have biological analogues. Organisms are constantly exchanging the materials that compose them for new materials, just as the ship of Theseus exchanged its old planks for new ones. Suppose, then, that I am an organism composed by the *x*s at time t_1. Gradually, the *x*s come to be replaced entirely by the *y*s at time t_2. Suppose, moreover, that at that time the *x*s themselves get reassembled into an exact replica of the organism composed by the *y*s. Which organism, if either, am I: the one composed of the *y*s, or the replica composed of the reassembled *x*s? Van Inwagen's answer is that I am the organism composed of the *y*s. The reason is that organisms persist on account of their lives, and the same life can continue in different physical materials:

> If the activity of the *x*s at t_1 constitutes or results from a life, and the activity of the *y*s at t_2 constitutes a life or results from a life, then the organism the *x*s compose at t_1 is the organism the *y*s compose at t_2 if and only if the life that the activity of the *x*s at t_1 constitutes or results from is the life that the activity of the *y*s at t_2 constitutes or results from. (1990: 148–9)

So long as the life constituted by the activity of the *x*s persists, I persist. Since my life at t_1, the life constituted by the activity of the *x*s at t_1, is the same as the life constituted by the activity of the *y*s at t_2, it follows that I am the organism composed of the *y*s at t_2.

The biological analogue of the statue and the lump argues that I share all the same parts with the mass of matter that composes me at exactly the same time. But two things cannot share all the same parts at exactly the same time, so I must be identical to the mass of matter. But the mass and I cannot be identical since we have different properties. The mass of matter existed (scattered throughout the biosphere) before I did, and unlike me, it can survive being squashed. Since the mass and I have different properties, it follows from Leibniz's law that we must be distinct. There is thus good reason to think both that the mass and I are identical and that the mass and I are distinct. Van Inwagen's (1990: 144) solution is to deny that there are masses of matter (or lumps, chunks, or hunks of matter). The particles that occupy my location compose nothing other than me, and in that case, if there is nothing—no mass of matter—located exactly where I am in addition to myself, there can be no question whether I am identical to a mass of matter.[5]

[5] It should be evident that van Inwagen rejects Composition as Identity, the thesis defended by David Lewis (1991) that an individual whole is in some sense identical to its parts. There are several well-rehearsed arguments against Composition as Identity. One is that the thesis is false at best and nonsensical at worst (Oliver 1994: 221–2; van Inwagen 1994: 210–13; Merricks 2001a: 21). Trenton Merricks states the point as follows: "One good reason to reject composition as identity is that it implies . . . that one thing (e.g. a whole) can be identical to many things (e.g. the whole's parts). But . . . one of the most obvious facts about *identity* is that while it holds both one-one (John is identical to Mr Smith) and perhaps even many-many (John and Mary are identical with Mr Smith and Ms Jones), it never holds one-many" (2001a: 21). In addition, Composition as Identity appears to imply a highly implausible commitment to mereological essentialism, the claim that a thing has all of its parts essentially (Merricks 2001a: 21–8). Roughly, since identity is necessary, if I am identical to my parts, $a_1, a_2, \ldots, a_n$, then I cannot exist without them, nor can they exist without me. Hence, any possible world in which $a_1, a_2, \ldots, a_n$ exist is a world in which I exist, and

Why does van Inwagen stop with organisms? Why not deny that all composites exist? Van Inwagen's most compelling answer, one developed more fully by Trenton Merricks (2001a), is that organisms have nonredundant causal powers that other alleged composites lack. The activities attributed to artifacts and natural bodies can be understood as disguised cooperative activities performed by simples, says van Inwagen. The chair, the mountain, and the planet don't do anything that cannot be exhaustively described and explained by appeal to the activities of mereological simples, but according to van Inwagen, not all activities are like this. Organisms are capable of doing things that cannot be done by simples alone but only by composite individuals. His example is thinking:

> Things can work together to produce light . . . But things that work together to produce light are not forced, by the very nature of the task set them, to produce light *by* composing a single object that emits light. And things that work together to support weight are not forced, by the very nature of the task set them, to support weight *by* composing a single object that holds things aloft . . . But things cannot work together to think . . . [P]lanning for tomorrow or feeling pain cannot be activities that a lot of simples can perform collectively, as simples can collectively shine or collectively support a weight . . . [T]hinking cannot be understood as a disguised cooperative activity. But *all* the activities apparently carried out by shelves and stars and other artifacts and natural bodies can be understood as disguised cooperative activities [performed by simples]. And, therefore, we are not forced to grant existence to *any* artifacts or natural bodies. (1990: 118, 122)

If van Inwagen is right, then we are forced to grant existence to ourselves since we engage in an activity, namely thinking, that cannot be performed by simples alone. Yet, if we grant existence to ourselves, says van Inwagen, then we should not deny it to other organisms. There are thus good reasons, van Inwagen thinks, for accepting that organisms exist while yet denying that artifacts and natural bodies do.

Van Inwagen's view implies a kind of property pluralism. Living things have properties, such as thinking, which are different from any of the properties had by simples. They cannot be understood as cooperative activities of those simples, but must be understood as properties unique to living wholes.

vice versa. It does not matter where $a_1, a_2, \ldots, a_n$ are located, or how they are arranged. If I am identical to $a_1, a_2, \ldots, a_n$, then I exist exactly if and where they do. Moreover, it is impossible for me to gain or lose any parts. This implies that I cannot be an organism, at least not as organisms are typically understood, since organisms are continuously exchanging materials with their environments. Both of these implications seem highly implausible. Van Inwagen (1994: 216–19) identifies a weaker version of the thesis, according to which composition is not really identity after all, it is merely analogous to identity. Here is how Lewis states the position: "Mereological relations . . . are strikingly analogous to ordinary identity. So striking is this analogy that it is appropriate to mark it by speaking of mereological relations—the many-one relation of composition, the one-one relations of part to whole and overlap—as kinds of identity. Ordinary identity is the special limiting case of identity in the broadest sense" (1991: 83). The problem with this version of the thesis, van Inwagen argues, is that opponents of Composition as Identity are not likely to find the notions of composition and identity to be analogous in the ways Lewis describes, and as a result, Lewis cannot use the Composition as Identity thesis to do the work he wants it to do, namely defending his commitment to unrestricted composition.

Let the foregoing remarks suffice for an overview of van Inwagen's account of composition. It is worth pointing out that van Inwagen's lives play precisely the kinds of theoretical roles that individual-making structures are supposed to play on the hylomorphic view. *Lives matter* on van Inwagen's view; they are ontological principles: whether the *x*s constitute a life makes a difference to whether or not some composite individual exists. Likewise, *lives make a difference*; they are explanatory principles: living things are capable of doing things that cannot be exhaustively described and explained using the conceptual resources used to describe and explain the materials that compose them (1990: 122, 180). Finally, *lives count* on van Inwagen's view; they operate as principles of unity (121) and persistence (145, 148): what binds the simples that compose me into a single being is that their activity constitutes a life, and what enables me to persist through changes in those simples is the persistence of that life. Because van Inwagen's lives play these roles, it is easy to use his view of composition as a basis for developing the hylomorphic view.

6.3 Hylomorphic Composition

The hylomorphic view of composition is similar to van Inwagen's in its outlines. Like van Inwagen's view, it provides a moderate answer to the SCQ. Configuring materials and being composed of materials are co-foundational concepts on the hylomorphic view, just as having a life and being composed of simples are co-foundational concepts on van Inwagen's. Likewise, just as van Inwagen restricts composition to living things, hylomorphists restrict it to structured individuals in general. According to hylomorphists, composition occurs when and only when an individual configures materials: there is a *y* such that the *x*s compose *y* if and only if *y* is an individual that configures the *x*s. I will call individuals that configure the materials composing them *structured individuals*.

Structured individuals are *emergent* individuals on the hylomorphic view.[6] There are empirically describable conditions that are sufficient to bring into existence new structured individuals with distinctive powers where previously no individuals with those powers existed. Principal among the powers of such individuals are powers to configure materials—powers that structured individuals are essentially and continuously manifesting (Section 6.1). A structured individual comes into existence exactly when its activity of configuring materials commences, and the materials it configures are precisely those that compose it. Structured individuals are thus emergent individuals who are essentially engaged in the activity of configuring the materials that compose them.[7]

[6] O'Connor and Jacobs (2003) argue that any view committed to emergent properties must posit emergent individuals as well.

[7] The idea that structured individuals are emergent is perhaps the most surprising element of the hylomorphic theory I defend. Consider, by contrast, a hylomorphic theory like Mark Johnston's (2006),

What exactly are activities? Do they represent a new ontological category in addition to those I've already discussed: individuals and properties? The answer is no: individuals engaging in activities are just individuals having properties of a specific sort, a sort I'm calling 'activities', although hylomorphists needn't insist on this term. Someone is free to say that my configuring the materials that compose me is not an activity at all, except in an extended sense. Hylomorphists needn't balk at this suggestion. They are free to adopt an alternative terminology. The important point is that configuring the materials that compose me, imposing my individual-making structure on them, is something I do.[8] There were empirically specifiable conditions that were responsible for bringing me into existence at a time, the very time at which I began configuring materials human-wise.[9]

When it comes to characterizing configuring activities, hylomorphists can adopt most of what van Inwagen says about lives, at least when it comes to the configuring activities of living things, the paradigmatic structured individuals. My life is identical to my configuring various fundamental physical materials at various times—an event that is self-maintaining, well-individuated, and jealous, and that has many other characteristics it is the business of the biological sciences to describe. An individual living thing does not configure exactly the same materials for very long since those materials are in constant flux, yet despite this, the individual maintains itself one and the same through all the changes on account of its ongoing configuring activity. That

which takes structures (what he calls 'principles of unity') to be relations among the material constituents of things. According to Johnston, what it is for a structured individual to be is for such-and-such parts to have the property or stand in the relation *R*, where *R* is the thing's principle of unity (2006: 658). For example, what it is for a human organism to be is for its parts to be related human-wise. The hylomorphic view I defend does not deny that there are many complex relations among the parts that compose a structured individual. Nor does it deny that some of those relations are both necessary and sufficient for a structured individual to exist. What it instead denies is the stronger canonical formulation Johnston proposes for the real essences of structured wholes: '*What it is* for a structured individual to be is…' The hylomorphic theory I defend proposes instead that what it is for a structured individual to be is for there to be an individual that configures its parts a certain way. For example, what it is for a human organism to be is for there to be an individual that configures its parts human-wise. The human-wise relation among a human's parts is necessary for it to exist, and it is also sufficient since the existence of structured wholes is both necessitated by and supervenient upon the conditions of their parts (Chapter 9). A human organism cannot exist unless its parts are spatially arranged human-wise, and necessarily if some physical materials come to be spatially arranged human-wise, a human organism will come to exist (or perhaps an exact replica of a human organism will come to exist if being a human requires having come into existence as the result of some type of evolutionary history). But on the hylomorphic view I defend, these claims needn't be taken to amount to a statement of a human's real essence. Some readers might consider this to be a finer point that is of little interest outside hylomorphic circles. If so, they are free to ignore it and move on.

[8] Marmodoro (2013) appears to endorse a similar view: she characterizes substantial form (what I have been calling 'individual-making structure') as an operation.

[9] Robert Koons appears to endorse a similar view: "Formal and material causation are, on my view, both real, diachronic causal connections… A composite substance exists at time *t* because its material components participated in an appropriately formal process in some interval of time immediately before *t*" (2014: 159).

activity is what unifies various materials into a single individual, both synchronically and diachronically, just as lives do on van Inwagen's account.

The hylomorphic view is also committed to a meta-mereology similar to van Inwagen's, one expressed by its commitment to ontological naturalism (Section 1.2). Hylomorphists take our best empirical descriptions, explanations, and methods as their starting point for ontological commitment. The hylomorphic view also has many of the same implications as van Inwagen's. It too is committed to a version of the Denial. It rejects the existence of artifacts, and masses (or chunks, hunks, or lumps) of matter, and can therefore solve puzzles like the statue and lump, the Ship of Theseus, and their biological analogues in the same way van Inwagen does.

The hylomorphic view is also committed to property pluralism. It implies that structured individuals have properties of at least two sorts: properties due to their individual-making structures (or their integration into individuals with individual-making structures), and properties due to their materials alone, independent of the ways they are structured. Subatomic particles, atoms, and molecules have physical properties such as mass irrespective of their surroundings. Under the right conditions, however, they can contribute to the activities of living things. Nucleic acids, hormones, and neural transmitters are examples. They are genes, growth factors, and metabolic and behavioral regulators. Each admits of two types of descriptions. They can be described in terms of the contributions they make to a structured system, but they are also independently describable in noncontribution-oriented terms. Descriptions of the former sort express the properties characteristic of structured individuals such as organisms and their parts. Descriptions of the latter sort express the properties they possess independent of their integration into structured wholes. A strand of DNA might always have various atomic or fundamental physical properties regardless of its environment, but it acquires new properties when it is integrated into a cell and begins making contributions to the cell's activities. It becomes a gene, a part of the cell that plays a role in, say, protein synthesis. Philosophers and biologists, such as Neil Campbell, who was quoted in Chapter 1, sometimes call the 'new' properties acquired by structured systems *emergent properties*. They have the following characteristics:

(1) They are first-order properties, not higher-order ones.
(2) They are not epiphenomenal, but make distinctive causal or explanatory contributions to the behavior of the individuals having them.
(3) They are possessed by an individual on account of its organization or structure.

Three remarks are in order about emergent properties. First, notice that there is no condition stipulating that emergent properties are in any sense produced or generated by lower-level materials, processes, or states. Although emergentists and epiphenomenalists typically add such a condition, hylomorphists do not. As we will see in detail when discussing the problem of emergence in Section 13.2, hylomorphists deny that emergent properties are in any sense causal byproducts of lower-level occurrences.

Second, Condition 1 is what distinguishes the hylomorphic view of properties from the view endorsed by many nonreductive physicalists. Nonreductive physicalists often claim that properties such as being in pain are higher-order properties, properties whose definitions quantify over other properties (Section 5.3). Being in pain, for instance, might amount to having some property whose instances satisfy a condition such as having a range of typical causes and effects (Section 11.4). According to Condition 1, logical constructions of this sort do not count as emergent properties. When we describe the contributions that, say, a DNA molecule makes to a cell's overall activity, we are not describing fundamental physical processes using an abstract vocabulary; we are rather describing first-order properties and relations that are distinctive of biological organization.

Third, Condition 3 is what distinguishes the hylomorphic view of properties from views like panpsychism. It implies that emergent properties are not aggregative properties such as mass. Eleanor has a mass of 50kg for the same reason a 50kg rock does: their fundamental physical constituents have masses that when taken in combination add up to 50kg. Condition 3 implies that emergent properties are not aggregative properties of this sort. In addition, Condition 3 implies that emergent properties are not possessed by any of an individual's parts since the individual-making structure responsible for them is not a feature of any one part considered in isolation; it is instead a feature of the whole those parts compose.

Despite the foregoing similarities between the hylomorphic view of composition and van Inwagen's, there are three noteworthy differences. First, van Inwagen is a committed atomist: he claims that fundamental physical materials are particulate. Hylomorphists are not committed to this. Nor are they committed to Aristotle's claim that fundamental physical materials are nonparticulate stuffs. The hylomorphic view defended here assumes that fundamental physical entities are capable of composing structured wholes, but it does not take a further stand on the natures of those entities. It instead leaves it to the relevant empirical disciplines to tell us what their natures are. In order to express this neutrality toward the nature of fundamental physical entities, I've used the term 'materials' to refer to them. In ordinary English, the term 'materials' can be applied both to discrete individuals and to continuous stuffs, something the term 'building materials' illustrates. It can be applied both to individual timbers and nails, and to stuffs such as glue, metal, and wood. When I use expressions like 'fundamental physical materials,' they should be taken to imply nothing specific about the particulate or continuous natures of those materials. According to hylomorphists, discovering what those natures are is a job for physicists.

Second, van Inwagen limits composition to living things. According to hylomorphists, living things are the paradigmatic structured individuals, but hylomorphists do not rule out a priori the possibility that there might be structured individuals of other sorts. Consider atoms and molecules. Van Inwagen claims that they do not exist; they are what he calls "virtual objects" which are "virtually composed" of

mereological simples (1990: 112). Hylomorphists are free to take the same stance as van Inwagen, but two more options are available to them. They can claim that atoms and molecules are structured wholes in their own right, distinct from living things and their parts, or they can claim that atoms and molecules are parts of living things and that the atom- and molecule-wise arrangements of fundamental physical materials we find outside organisms are virtual objects in van Inwagen's sense. Which stance they take depends on broadly empirical considerations. Roughly, if atoms and molecules have powers distinct from those which can be exhaustively described and explained by appeal to fundamental physical materials alone, then there are grounds for claiming that they are not mere aggregates of fundamental physical materials, but are distinctive individuals in their own right (Section 7.4).

This difference between van Inwagen's view and the hylomorphic one means that hylomorphism is compatible with two different pictures of how structured individuals emerged on the cosmic scene. According to the first picture, which is like van Inwagen's, the only structured individuals are living things. Prior to the advent of life, the universe contained only fundamental physical materials that were spatially arranged atom-wise or molecule-wise, but those spatial arrangements of physical materials composed nothing. According to the second picture, which claims, unlike van Inwagen, that there are structured individuals other than living things, some spatial arrangements of physical materials, perhaps atoms or molecules, or something else, are not merely quantities of fundamental physical materials standing in certain spatial arrangements; they instead compose individuals in their own right with individual-making structures that distinguish them from unstructured physical materials. The first picture takes the emergence of life to mark a sharp ontological break in the natural world—the first time in the universe's history that there existed anything but fundamental physical materials. The second picture sees the emergence of life as part of a more gradual process in which the universe produced by stages increasingly complex structured individuals. Which picture is accurate? Hylomorphists take this to be an empirical question, one that can only be answered by determining which individuals in the world have powers distinct from those which can be described and explained exhaustively by appeal to fundamental physical materials alone.

I'll say more about hylomorphists' stance on the issue of whether there are nonliving structured individuals in Section 7.4, but for the time being, the important point is that hylomorphists countenance a larger ontology of material beings than van Inwagen. His ontology of material beings includes only mereological simples and organisms. Hylomorphists' ontology includes fundamental physical materials, structured individuals, and their parts. Moreover, since hylomorphists take it to be an empirical matter what structured individuals there are, they are open to the possibility that there might be structured individuals other than living things. Living things are nevertheless the paradigmatic structured individuals, and for that reason, I will often use examples and adopt expressions that have a biological orientation.

This biological focus should not be taken to imply, however, that what I say cannot be applied mutatis mutandis to structured individuals of nonbiological sorts if there are any.

A third difference between the hylomorphic view and van Inwagen's is that the hylomorphic view of parts is less revisionary than van Inwagen's. Van Inwagen (1981, 1990) denies that there are organic parts such as eyes, hearts, and kidneys; according to him, the only proper parts are fundamental physical particles and individual cells. Hylomorphists, by contrast, accept that there are parts such as eyes, hearts, and kidneys. Their reasons for doing so are again broadly empirical. Roughly, our best descriptions and explanations of human behavior postulate parts like these, and this gives us good reason to think such parts exist. It is worth considering the argument for this claim in greater detail.

There are multiple ways of dividing things into parts. A hammer can be exhaustively decomposed into functional parts such as the head, the claw, and the handle—parts that contribute to the hammer's overall operation. It can also be exhaustively decomposed into spatial parts such as the top third, the middle third, and the bottom third, or into the spatial parts obtained by dividing it along the lines of a three-dimensional coordinate grid with metric units. It can also be exhaustively decomposed on the basis of the materials that compose it: the metal part versus the wooden part. Different principles of part identity and individuation yield different inventories of parts, all of which might be consistent with principles of formal mereology, such as the transitivity and asymmetry of proper parthood.[10] Because there are many different principles for part identity and individuation, whenever we want to consider the parts that a particular individual has, we need to determine which principle is (or which principles are) best suited to our descriptive and explanatory interests. If those interests include understanding what that individual is and what it does, and we accept ontological naturalism (Section 1.2), then empirical adequacy is an important criterion for making that determination.

Ontological naturalism, recall, claims that when it comes to determining what exists, empirical investigation—paradigmatically science—is our best guide. Given reasonable assumptions, ontological naturalism has the following implication: when it comes to choosing principles of part identity and individuation for various individuals, we should choose principles that reflect our best empirical descriptions and explanations of their behavior. Roughly, if principle P does a better job enabling us to describe and explain the behavior of Ks than principle P^*, then we have good reason to accept that Ks have parts that are identified and individuated by P instead of P^*.

[10] The transitivity of parthood is the principle that if x is a part of y, and y is a part of z, then x is a part of z. The antisymmetry of parthood is the principle that if x is a part of y, and y is a part of x, then $x = y$ (the tacit assumption being that everything is part of itself).

Consider now living things such as human organisms. Like the hammer, a human can be divided along purely spatial lines into thirds, or fifths, or along the lines of a three-dimensional coordinate grid with metric units. But biologists, neuroscientists, and psychologists are more interested in dividing humans and other organisms along functional rather than purely spatial lines.[11] In addition, when they divide organisms into functional parts, they take those parts to stand in hierarchical relations. As Ernst Mayr puts it, "Organisms are many-level ordered systems" (1997: 3). Perhaps the best example of how the biological sciences divide organisms into hierarchically ordered functional parts is provided by the method of scientific investigation that philosophers have sometimes called *functional analysis* (other names include 'mechanical decomposition' or 'functional decomposition').[12]

Functional analysis is a method that biologists, cognitive scientists, engineers, and others frequently employ to understand how complex systems operate. It involves analyzing the activities of those systems into simpler subactivities performed by simpler subsystems (Fodor 1968b; Simon 1969; Wimsatt 1974; Cummins 1975; Dennett 1978; Lycan 1987: chapter 4; Bechtel and Richardson 1993; Haugeland 1998; Glennan 2002; Bechtel 2007, 2008; Craver 2007). Consider a complex human activity such as running. Functional analysis reveals that running involves, among other things, a circulatory subsystem that is responsible for supplying oxygenated blood to the muscles. Analysis of that subsystem reveals that it has a component responsible for pumping the blood—a heart. Analysis of the heart's pumping activity shows that it is composed of muscle tissues that undergo frequent contraction and relaxation, and these activities can be analyzed into the subactivities of various cells. Analyses of these subactivities reveal the operation of various organelles that compose the cell and that are composed in turn of complex molecules. The cell membrane, for instance, is composed of a double layer of phospholipids, each of which has a hydrophobic end which repels water, and a hydrophilic end which attracts it. Analysis of the water-attractive end reveals that it is composed of a phosphate group with a distribution of electrons capable of attracting water molecules. The electrons are able to do this because they are negatively charged. If they do not have their charges on account of the activities of some lower-level subsystems, but as an unanalyzable matter of fact, then no further functional analysis is possible.

[11] Carl Craver (2007: chapter 5) calls purely spatial parts 'pieces' and parts in the functional sense 'components.' He takes himself to be adopting a distinction drawn by Stuart Kauffman (1971: 260). John Heil (2003: 100) also suggests something like the distinction between merely spatial parts and parts of other sorts, which he calls 'substantial parts.' Mark Johnston (2006: 656–7) makes a similar distinction between what he calls 'genuine parts' and 'fiat parts,' and Kit Fine distinguishes spatial parts from parts that "correspond to a mereological division via the operations of variable and rigid embodiment" (1999: 72).

[12] The term 'functional analysis' is due to Cummins (1975). Becthel (2008) calls it 'mechanistic decomposition' or 'functional decomposition.' Craver (2007) subsumes it under the heading of 'mechanistic explanation.' He takes Cummins' notion of functional analysis to be the exemplar of what he calls the 'systems tradition,' but argues that Cummins fails to provide an adequate account of mechanistic explanation. For that reason he distances himself from the term 'functional analysis.'

Functional analysis is important because it provides a basis for understanding the kinds of parts postulated by descriptions and explanations in the biological sciences. Those parts are subsystems or components, things that contribute in empirically specifiable ways to the activities of the wholes to which they belong. Saying that *x* is a part of *y* implies that *y* engages in some activity which is capable of being functionally analyzed into subactivities, and that *x* performs one of those subactivities. To say that my heart is part of me is to say that my activities can be given an analysis into subactivities (or those into further sub-subactivities, and so on), and that one of those subactivities (or sub-subactivities or sub-sub-... subactivities) is performed by the heart. What distinguish different parts of me from each other, moreover, are the different ways they contribute to my activities: different parts contribute in different ways.

A principle of part identity and individuation along these lines has been suggested by several philosophers of biology, including William Bechtel, a philosopher of neuroscience.[13] According to Bechtel, something qualifies as a component part of a complex system—what he calls a *mechanism*—only if it performs an operation that contributes to the activity of a whole:

> The component parts of a mechanism are the entities that perform the operations which together realize the phenomenon of interest. A structure within the mechanism may be well delineated (it has boundaries, continues to exist over time, is differentiated from the things around it, etc.). However, if it does not perform an operation that contributes to the realization of the phenomenon, it is not a working part of that mechanism. For example, while the gyri and sulci of the brain are well delineated, they are not working parts of the brain but byproducts of the way brains fold to conserve the length of axons. (Bechtel 2007: 180)

Philosophers of biology and neuroscience, like Bechtel, have been attracted to a view of parts along these lines because this is the type of view suggested by actual work in biology and neuroscience—both the methods of those sciences and the kinds of explanations they employ (Craver 2007: chapter 5).

Two clarifications are in order about functional analysis. First, a remark about the name: 'functional analysis' is a name that has been used by some philosophers, but other philosophers and many biologists call the method 'reduction' (Campbell et al. 1999: 4). This notion of reduction is different from the notion typically discussed in connection with mind-body problems (Jaworski 2011: 277). Reduction in the latter sense typically concerns the ability of one conceptual framework to take over the descriptive and explanatory roles of another (a point discussed in greater detail in Section 11.3). To claim that, say, psychology is reducible to neuroscience implies that it is possible in principle for neuroscience to take over all the descriptive and explanatory roles that psychology currently plays. Any description or explanation we normally express in psychological terms can be rewritten in neuroscientific ones.

[13] Carl Craver (2007: chapter 5) also endorses this type of view.

Ernest Nagel's (1961) influential account of reduction formulated this idea by appeal to bridge principles—empirically warranted premises that identified the entities postulated by the reduced theory with entities postulated by the reducing one.[14] These theoretical identifications would enable us to rewrite the descriptions and explanations of the reduced theory as descriptions and explanations of the reducing theory. If pain were identical to brain state A, and anxiety were identical to brain state B, then the statement 'pain triggers anxiety' could be rewritten as 'Brain state A triggers brain state B.' If psychological descriptions and explanations could be rewritten as neuroscientific ones in this way, then neuroscience would be capable of taking over all the descriptive and explanatory roles psychology plays. Psychology would be reducible to neuroscience.

By contrast, when biologists speak of reduction, they are typically not speaking of the relation between conceptual frameworks I've just described, but of a *method* for studying complex systems—the method I've been calling 'functional analysis.' Here is an example taken from the biology textbook quoted earlier:

> Reductionism—reducing complex systems to simpler components that are more manageable to study—is a powerful strategy in biology . . . Biology balances the reductionist strategy with the longer-range objective of understanding how the parts of cells, organisms, and higher levels of order, such as ecosystems, are functionally integrated. (Campbell et al. 1999: 4)

The authors clearly have in mind what they call a research *strategy*—a method for studying complex things. A commitment to employing this method does not imply a commitment to reduction in the aforementioned philosophical sense.[15] It might be impossible for neuroscience to take over the descriptive and explanatory roles of psychological discourse even though it is possible and even necessary to use functional analysis to understand how humans can engage in psychological activities. In fact, this is precisely what hylomorphists claim.

According to hylomorphists, explanations of living behavior are not reducible to descriptions of the lower-level mechanisms revealed by functional analysis because of the distinctive explanatory contributions a living thing's biological structures make.

[14] Nagel's original account did not require that bridge principles be identity statements, but many critics argued forcefully to the contrary, including Sklar (1967), Schaffner (1967), and Causey (1977: chapter 4). See Jaworski (2011: 123–4) for further discussion of this point.

[15] John Bickle (2003) refers to the philosophical sense of reduction as 'ruthless reduction' (although he denies that ruthless reduction requires bridge principles that take the form of identity statements); Bechtel (2007: 173–4) and others have distinguished this from reduction in the methodological sense—what we are calling 'functional analysis.' David Chalmers too argues that a commitment to using functional analysis does not imply a commitment to reduction in the sense I've described. He calls the explanations yielded by functional analysis 'reductive explanations': "A reductive explanation of a phenomenon need not require a reduction of that phenomenon . . . [P]henomena [such as learning] that can be realized in many different physical substrates . . . might not be reducible in that we cannot *identify* learning with any specific lower-level phenomenon . . . Reductive explanation of a phenomenon should . . . not be confused with a reduction of a high-level *theory*" (1996: 43). To avoid confusion, I prefer Bechtel's term for referring to the kinds of explanations yielded by functional analysis: 'mechanistic explanations.'

Why should we believe that hylomorphists are right about higher-level structures making explanatory contributions beyond the contributions made by lower-level things? This is a topic discussed in greater detail in Chapters 10–14. Here I'll merely sketch the hylomorphic answer.

Briefly, hylomorphists insist that irreducibility is supported by empirical considerations. As an empirical matter of fact, they say, higher-level structural discourse provides effective descriptions and explanations independent of any claims about reducibility. Consider Bechtel's observations about descriptions and explanations in psychology and other special sciences:

> [The] mechanistic explanations [provided by functional analysis] are in fact compatible with a robust sense of autonomy for psychology and other special sciences . . . In virtue of being organized systems, mechanisms do things beyond what their components do . . . Organization itself is not something inherent in the parts. Accordingly, investigators who already understand in detail how the parts behave are often surprised by what happens when they are organized in particular ways . . . [T]he organization of the components typically integrates them into an entity that has an identity of its own. As a result, organized mechanisms become the focus of relatively autonomous disciplines . . . This autonomy maintains that psychology and other special sciences study phenomena that are outside the scope of more basic sciences but which determine the conditions under which lower-level components interact. In contrast, the lower-level inquiries focus on how the components of mechanisms operate when in those conditions . . . The fact that mechanisms perform different activities than do their parts manifests itself in the fact that the activities of whole mechanisms are typically described in different vocabulary [sic] than are component operations. Traditional accounts of theory reduction implicitly recognized this fact by requiring bridge principles to connect the different vocabularies used in different sciences, but little notice was given as to why different sciences employ different vocabularies. The vocabulary used in each science describes different types of entities and different operations—one describes the parts and what they do, whereas another describes the whole system and what it does. (2007: 174, 185–6)

If Bechtel's observations of scientific practice are correct, higher-level empirical disciplines and lower-level ones have different subject matters on account of the ways things are organized or structured. Because higher- and lower-level disciplines deal with different subject matters, they have different vocabularies, and provide different kinds of explanations, and these different vocabularies and explanations make higher-level disciplines autonomous—irreducible to lower-level disciplines in the philosophical sense I've described. In light of these kinds of observations about the autonomy of higher-level sciences, and the role structure or organization plays in explaining it, hylomorphists insist that the burden of proof is on their opponents to establish that claims about biological, psychological, or social structures are reducible to claims about things that lack them.

A second note about functional analysis: the notion of function that gives functional analysis its name is different from the notion of function discussed in connection with functionalism in the philosophy of mind. According to classic functionalist

theories of mind, mental states are postulates of abstract descriptions framed in terms analogous to those used in computer science—descriptions that ignore the physical details of a system, and focus simply on inputs to it, outputs from it, and internal states that correlate the two (Putnam 1975b: essays 18–21; Jaworski 2011: 136–41). When it comes to functional analysis, by contrast, the notion of a function is not abstract in this way, and it has a teleological dimension: subsystems contribute to the activities of the wholes to which they belong, and that contribution is their reason or "purpose" for belonging to the system (Lycan 1987; Sober 1985): the purpose of the spark plug is to ignite the fuel, the purpose of the heart is to pump the blood, and so on.

Teleological functionalism is a type of functionalist theory that appeals to a teleological notion of function along these lines as well (Jaworski 2011: 153–4). Lycan's (1987) homunctionalism is an example. Like functionalist theories of all sorts, however, teleological functionalism claims that higher-level discourse is abstract discourse: higher-level properties are higher-*order* properties—logical constructions that quantify over lower-order properties. We have already seen that our preferred ontology has no place for higher-order properties (Section 5.3). According to hylomorphists, higher-level properties are first-order properties in their own right. Consequently, although teleological functionalists and hylomorphists both claim that a system's components contribute teleologically to its overall operation, they disagree about how the notion of contribution is to be understood. Teleological functionalists claim that descriptions of higher-level phenomena are simply abstract descriptions of lower-level occurrences. Hylomorphists claim that higher-level descriptions correspond to distinctive natural structures that factor into descriptions and explanations of living behavior in ways that cannot be eliminated, reduced to, or paraphrased in favor of lower-level descriptions and explanations.

6.4 Biofunctional Parts

Not just any kind of contribution to an organism's activities can qualify something as a part of the organism on the hylomorphic view. The computer, the pencil, and the cell phone all contribute to my activities, but they are not biologically configured by me; they are not caught up in my life in van Inwagen's sense. It will be helpful, therefore, to say a bit more about what the relevant kind of contribution is.

To be caught up in something's life in van Inwagen's sense is, intuitively, to belong to its living tissue. What this amounts to is ultimately an empirical matter, something it is the business of the biological sciences to describe. But it might be possible to sketch the rough outlines of an account based on the idea that living things are composed of cells. Most biologists would agree that cells are the simplest systems that display the range of behaviors that characterize living things: the acquisition and use of nutrients from the environment, the excretion of waste materials, reproduction, response to environmental stimuli, and so on. Imagine, then, that a functional

analysis of Gabriel's running reveals a circulatory system with a subsystem that pumps blood, a human heart. Functional analysis of the heart's pumping activity reveals that it is composed of muscle tissues that are composed in turn of cells. Because cells are living things, something that is composed of cells is composed of living materials. So because Gabriel's heart is composed of living materials, his heart qualifies as a part of him. But being composed of cells at some level is not the only way that something can be caught up in the life of an organism. The organelles of single-celled organisms such as amoebas are parts of those organisms; the organisms configure them, they are caught up in those organisms' lives, yet they are not composed of cells. What qualifies them as parts is not that they are composed at some level of cells, but that they contribute in some way to the activities of cells.[16] A cell membrane, for instance, is caught up in the life of the cell because it contributes to the cell's homeostatic activity, the maintenance of its internal environment. The phospholipids that compose the membrane also qualify as parts of the cell since they contribute to the activity of the cell membrane. Likewise, the electrons at the hydrophilic end of the phospholipid molecule qualify as parts of the cell because they contribute to the molecule's water-attracting activity.

In short, then, being caught up in the life of an organism can be understood to consist either in composing cells or in being composed of cells. My heart qualifies as a part of me because it is composed of cells, whereas the electrons in my heart qualify as parts of me because they compose cells. Let us call this account of parthood with its specifically biological orientation a *biofunctional account of parthood*, and call parts that are identified and individuated by the contributions they make to the activities of living things *biofunctional parts*.[17]

A biofunctional account of parthood has several noteworthy implications. For one thing, it implies that there are no artificial organs. So-called "artificial organs"—artificial hearts, pacemakers, prosthetic limbs, cochlear implants, neural stimulation devices, and so on—are not parts of the organisms in which they are implanted. Imagine again a functional analysis of Gabriel's running, but this time, imagine that the blood-pumping component of Gabriel's circulatory system is not a human heart, but a different device. Functional analysis of its activities reveals that it does not have cells as components at any level. It is not composed of living materials, but of titanium and plastic. Since the device is neither composed of cells nor composes cells, it is not part of Gabriel's living tissue; he does not biologically configure it; it is not caught up in his life; it is merely an artifact that contributes to his activity in a way that

[16] An analogous point could be made of parts such as the vitreous humor of the eye, which is not composed of cells. It contributes to the activity not of a cell, but of an organ.

[17] I mentioned in Section 6.3 that hylomorphism is compatible with the existence of structured individuals other than living things. If there are nonliving structured individuals, then it will be necessary to construct a nonbiological analogue of a biofunctional account of parthood. The parts of nonliving wholes will be things that contribute to the distinctive activities of those wholes in something analogous to the way biofunctional parts contribute to the activities of living wholes.

approximates the way a human heart would. Artificial devices of this sort can contribute to our activities in important ways—ways that are every bit as important to our survival as the biofunctional parts they are meant to replace. But that importance does not qualify them as biofunctional parts. They are merely artifacts, albeit organ-like artifacts that contribute to our activities in ways that approximate the ways biofunctional parts would.

In addition, a biofunctional account of parthood implies that there are no severed parts. So-called "severed parts" are not parts at all. Once a hand has been severed, the hand-shaped materials that remain no longer compose a hand. At best, they amount to a hand only in name, as Aristotle says.[18] Since they are no longer being configured by an organism, they do not compose anything; they are no longer caught up in a life.[19] In speaking of a severed hand, we are speaking of materials that used to contribute to the activity of an organism in the ways associated with hands. We might speak of the *remains* of a hand—tissues, cells, or other materials that used to compose a hand. We speak in a similar way of human remains or the remains of a deer. Sometimes we use singular terms such as 'corpse,' 'cadaver,' and 'carcass' to refer to something's remains, but strictly speaking, there are no such things on the hylomorphic view. When I die, the physical materials that remain do not compose a single object, a corpse or cadaver, nor is there a single object, a carcass, composed by the remains of the deer. To speak of a corpse or carcass is to speak of physical materials that used to compose an organism but that no longer do. They are materials that used to be caught up in a life, that used to be structured animal-wise, but that no longer are.

One question that arises at this point concerns the ontological status of parts on the hylomorphic view I defend. On the one hand, I've suggested that parts of structured wholes, such as electrons, are capable of existing independently of those wholes. On the other hand, I've just argued that parts of structured wholes are not capable of existing independently of the wholes to which they belong. How are we to reconcile these claims? The answer is that hylomorphists take there to be two kinds of parts. Some parts are capable of existing independent of the wholes to which they belong; that is, their status as parts is something they enjoy only contingently. An electron is capable of existing independent of a structured whole, but it is also capable of being part of one. Parts like the electron are often referred to using the expression 'the matter,' as in "The matter that composes me now is not the same as the matter that composed me seven years ago, nor will it be the same as the matter that composes me seven years hence" (Section 7.3). On the hylomorphic view, however, not all parts are like the electron. Some are like the heart: they are not capable of

[18] See, for instance, *On the Soul* 412b21–3; *Metaphysics* 1035b24; *Meteorology* 390a10–12; *On the Generation of Animals* 734b24–7.

[19] Van Inwagen denies the existence of so-called severed parts for precisely this reason. As we've seen, he also denies the existence of unsevered parts other than cells and mereological atoms.

existing independent of the wholes to which they belong. The reason, as we will see in Section 7.1, is that they enjoy a unity—a subordinate unity—that depends on the wholes to which they belong. The unifying activity that makes diverse materials compose a heart depends on the unifying activity of a more inclusive whole in a way that the unifying activity that makes an electron does not. Because of this, electrons but not hearts are capable of existing independent of more inclusive wholes. If my heart were removed from my chest cavity, it would cease to exist; the remaining physical materials would not compose a heart, they would simply be physical materials that used to compose a heart but that no longer compose anything. If those materials were placed back into me, or if they were transplanted into someone else, they would come to compose a heart once again. What is true of the heart is not true of an electron. If the electron that is part of me were removed, it would not cease to exist. Nor would a new electron come into existence if the original electron were reincorporated into my living tissue. What would happen is that the electron would gain new properties, for its reincorporation into my living tissue would consist in its making biofunctional contributions to my activities that it was not making before.

6.5 Ackrill's Puzzle and the Thomistic Theory of Parts

Two remarks are in order about the hylomorphic theory of parts I've described. First, it dovetails with an attractive solution to a well-known puzzle about the application of Aristotelian hylomorphism to living things which was articulated by John Ackrill (1972–3). On the one hand, a central feature of Aristotle's account of change is that matter has whatever form it does only contingently. The bronze that is shaped into a statue needn't be statue-shaped. On the other hand, Aristotle tells us that living bodies have their souls, which are their forms, essentially. A living body that has lost its soul, he tells us, is a body only in name—no more a body than the eye of a statue is an eye (*De Anima* II.1 412b10–24). The body of a living thing is the matter of its soul, which is its form. So Aristotle appears committed to saying—contrary to his general hylomorphic principles—that the matter of a living thing does not have its form only contingently since that matter (the body) cannot exist without the form that it actually has (the soul). Aristotle's hylomorphic framework thus appears to be inapplicable to living things, yet for Aristotle, living things are the paradigmatic examples of hylomorphic compounds.

One attractive solution to Ackrill's puzzle distinguishes two kinds of matter: matter that Aristotle sometimes characterizes as *organic* (*De Anima* II.1 412b1) and matter of a more basic sort which is the matter of organic matter (*Metaphysics* VIII.4 1044a15–25). The organic matter is the array of fully-formed bodily organs (for example, the brain, liver, heart, eyes, and limbs) that enable the living thing to engage in the activities that characterize it as a living thing of its particular kind. But each of those organs is composed of more basic materials, ultimately earth, air, fire, and water. So when Aristotle says that the living body has its soul essentially, he is

referring to the organic body which is the matter of the soul. If this body loses its soul, it ceases to exist. But a living thing comprises more than just this organic body as its matter. It comprises several kinds of matter, Aristotle tells us, since the organic body is composed of other materials (*Metaphysics* VIII.4 1044a20–2)—quantities of earth, air, fire, and water, let us suppose. These materials are only contingently ensouled: they have only contingently the form of a living thing. On this interpretation, Aristotle is saying that when a quantity of earth, air, fire, and water take on (contingently) the form of, say, a human being, there comes to be a structured whole, a body with organs that empower it to engage in the various activities which categorize it as a human. That organic body has the human form essentially; if that form were lost, the organic body would cease to exist. All that would remain would be the quantities of earth, air, fire, and water that used to compose it, but that no longer do.

The aforementioned distinction between two kinds of matter corresponds to the distinction I've drawn between two kinds of parts. There are, on the one hand, parts like the heart which cannot exist apart from more inclusive wholes, and hence which belong to such wholes essentially. These are parts corresponding to the organic matter described above. On the other hand, there are parts like the electron which can exist apart from more inclusive wholes, and hence which belong to those wholes only contingently. These are parts corresponding to the more basic kind of matter described above. The account of parts I've described thus dovetails with an attractive solution to Ackrill's puzzle.

Second, this view of parts differs from the view many Thomists endorse. The Thomists I have in mind claim that all parts are like the heart: they all depend for their existence on more inclusive wholes. If I inhale an oxygen atom, that atom does not survive being incorporated into me. It is instead replaced by something else—something that perhaps has many of the same characteristics as the original atom, but that is nevertheless numerically different from it. The resulting view is similar to Michael Burke's (1996) disappearing matter view.

In defense of this view, Thomists frequently appeal to the doctrine of the unicity of substantial form. Oderberg describes it as follows:

> Unicity of form means that for any one substance, there is one and only one substantial form which it possesses. This is because a substance is one kind of thing, and substantial form determines the kind of thing it is. Hence when a substance comes into being it does so by virtue of acquiring a single substantial form, and when it loses that form it ceases to exist altogether as that kind of thing, even if something else is left over which is not that kind of thing. (2007: 68)

The substantial forms Oderberg mentions correspond to individual-making structures in the hylomorphic framework I've described. Substantial forms *count* on the Thomistic view, just as individual-making structures count on the hylomorphic view I've articulated. Both are primitive principles of unity. Oderberg initially applies the

doctrine to kinds that are related as genus and species.[20] A dog does not possess multiple substantial forms, one by virtue of which it is a dog, and another by virtue of which it is an animal. There is but a single substantial form which is responsible for it being both a dog and an animal. This is a plausible result, which accords with the hylomorphic view I've articulated. The latter view claims that properties, including structures, are all fully determinate (Section 3.3). The predicate 'is an animal' is a determinable predicate; it does not express a property in its own right, so it cannot express a determinable individual-making structure; there are no such structures.

But Oderberg and many other Thomists take the doctrine of unicity to imply more than this; they take it to imply that an oxygen atom cannot be incorporated into my living tissue, that it can only be replaced by something else. Oderberg explains:

> [S]ubstantial form permeates the entirety of the substance that possesses it . . . there is as much dogginess in Fido's nose and tail as in Fido as a whole . . . down to the very chemical elements that constitute Fido's living flesh. To use the Scholastic terminology, the chemical elements exist *virtually* in Fido, not as compounds in their own right but as elements fully harnessed to the operations of the organisms in which they exist . . . Suppos[e] there to be elementary particles . . . and suppose[e] these to be quarks . . . [I]n the existing substance the quarks *have* no substantial identity of their own, their behavior having been fully yoked to the function and operations of the substance in which they exist . . . the quark is ontologically dependent on the whole of which it is a part, but its causal powers persist, albeit in a way radically limited by the whole. (2007: 70–1)

What is difficult to appreciate here is how it follows from the claim that there is only one substantial form per individual, that quarks, leptons, oxygen atoms, or what have you cease to exist when they are incorporated into a structured whole.

To bring the point into focus, let us formulate the doctrine of unicity in terms of individual-making structures as follows:

> *Thomistic Unicity*: For any composite individual, x, with an individual-making structure, S, which unifies fundamental physical materials in such a way that they compose x at t, there is no other individual-making structure, S^*, which unifies those same materials in such a way that they compose a different kind of individual at t.

Thomistic unicity rules out the possibility that the very same quantity of physical material might compose both a human, say, and a different kind of individual such as a chimp or a large dog. Located exactly where I am, there is at most one (fully determinate) individual-making structure. What Thomistic unicity doesn't appear to

[20] This is apparently how the principle was originally understood in the medieval debate about unicity (Callus 1961).

rule out, however, is the view of parts that I've articulated. That view is committed to the following thesis:

> *Hylomorphic Plurality*: It is possible for there to be a composite individual, x, with an individual-making structure, S, which unifies fundamental physical materials in such a way that they compose x at t, and another individual-making structure, S^*, which unifies a different quantity of fundamental physical materials in such a way that they compose a part of x at t.

For example, I have an individual-making structure that unifies physical materials in such a way that they compose me at t, but the oxygen atom in my lung has an individual-making structure that unifies a different quantity of physical materials in such a way that they compose it. Granted, the materials that compose the oxygen atom are included among the materials that compose me as a whole, but it is difficult to see how this should make a difference. Oderberg suggests that according to Thomists, the reason a thesis like hylomorphic plurality is ruled out is that the behavior of the physical materials that would ostensibly compose the oxygen atom have been "fully yoked" to the activities of the more inclusive whole to which they belong. But why suppose that the only way in which the behavior of those materials could be fully yoked to the biological imperatives of a whole should require that the oxygen atom they originally composed cease to exist? To make matters still more puzzling, Oderberg says that the causal powers of the original object still persist, that their operation is simply limited by the whole. What remains unclear is why this should lead us to insist that the original oxygen atom must cease to exist. Why can't the powers of the original atom be limited by the whole to which it belongs? Why must it be obliterated and replaced by a proxy?

The alternative hylomorphic view I've defended claims that the oxygen atom and the materials composing it continue to exist when incorporated into a more inclusive whole along with all the causal powers they originally possessed. The whole might nevertheless have other parts whose powers inhibit the manifestation of the powers of the oxygen atom, or parts whose powers are reciprocal with the powers of the oxygen atom, and because of that, the whole is able to limit or direct the way in which the oxygen atom manifests its powers. It is able to harness the powers of the oxygen atom in ways that contribute to its own biological ends. By analogy, a brick retains all of its powers when incorporated into a larger thing such as a house. It doesn't lose those powers, nor does it cease to exist when incorporated into the larger whole. Rather, the whole comprises other parts besides the brick, and the powers of these other parts, in combination with those of the brick, enable the builders to achieve their architectural ends. What rules out this picture? I cannot discern anything in Oderberg's comments that answers this question.

Anna Marmodoro suggests one possible answer: there might be no way of explaining how individual-making structures manage to unify diverse materials unless one endorses the Thomistic view:

How does the substantial form unify the elements in a substance[?] . . . [T]he unification of the parts of a substance into one is not achieved by any item that relates them . . . [T]he substantial form that unifies the elements of a substance is a *principle*. Since what is needed is the shedding of only the distinctness of the elements, the role of this unifying principle must be just that . . . [T]he substantial form . . . is *an operation* on the elements of a substance, stripping them of their distinctness . . . The unification of elements is achieved through their *re-identification* in terms of the role allotted to them by the substantial form. The elements are arithmetically and even qualitatively diverse, but are unified into a single entity by . . . *changing, i.e. transforming into something different*. (2013: 17)

According to both the Thomistic view and the hylomorphic view I've described, structures are primitive principles of unity. The structured whole configures the materials that compose it, and that configuring activity unites into one the diverse materials that compose the whole. What is it about that activity that makes the diverse materials one? According to the Thomist, they are made one by ceasing to be diverse; what enters into a composite individual ceases to exist and is replaced either by a part of the whole (which depends on the whole for its existence) or by the whole itself.

Marmodoro sometimes suggests the latter view; the parts of a composite whole, she says, exist "only potentially" (2013: 15). Koons (2014) calls this *Aristotelian Parts-Nihilism*. Why do the parts of a whole still seem to be diverse on this view? Perhaps the Thomist will say that any distinctions we draw among diverse parts of a unified whole are purely conceptual, that we choose to view the whole through the lens of this or that principle of part identity and individuation and come thereby to posit diverse parts. Whatever the case may be, a Thomist who endorses Parts-Nihilism denies that there are in fact any parts within the whole. Koons argues that this view faces serious difficulties. Among other things, it makes it difficult for the Thomist to account for internal locomotion (2014: 165), and may force the Thomist to treat regions of space, instead of individual substances, as bearers of causal powers (2014: 161–2).

The more charitable interpretation of Marmodoro's view is the former: what enters into a composite individual ceases to exist and is replaced by a part of the whole that depends on the whole for its existence. The problem with this view, as Koons sees it, is that it requires Thomists to deny a basic hylomorphic tenet: the Aristotelian idea that in any process of change or becoming, something exists, one and the same, through that process. "It is not enough," he says, "for there to be (before and after a case of substantial change) things that are quantitatively and qualitatively similar to each other, even exactly similar. There must be some one thing that endures through the change as its ultimate subject" (163). Why must hylomorphists endorse this idea? If they do not, Koons argues, they end up denying the very existence of change in favor of something like temporal parts theory:

Suppose that there could be a change with no enduring subject. If such a thing could happen somewhere at some time, it could happen everywhere at all times . . . [W]hat is possible in a given

situation cannot depend on what actually happens in remote situations . . . But a world in which there are never any enduring subjects of change is nothing more than a four-dimensional block of qualities—the sort of static block universe . . . endorsed by four-dimensionalists and Neo-Humeans. From an Aristotelian perspective, such a world would lack any real change or time at all. Hence, the very idea of substrateless change is incoherent. (2014: 163)

Koons suggests that a Thomist could respond by positing quantity tropes that persist through an instance of substantial change. But this response presents Thomists with a dilemma: either they must deny that tropes are nontransferable, something that poses difficulties for trope theorists (Section 3.2), or else they must posit persisting bearers for the persisting quantity tropes. But it is difficult to see what the persisting bearers could be other than regions of space, and given the account of properties defended in Chapters 2–5, that would appear to commit Thomists to attributing powers to regions of space instead of individual substances—a very un-Aristotelian position for Thomists to take.

There are other difficulties as well. What Dean Zimmerman says about Michael Burke's disappearing matter view applies to the Thomistic view as well:

[The] picture has its problems. If the matter now constituting my body is not the same as the matter that ceased to be as it came to constitute me, what is it and where did it come from? Either there is really no such thing as the matter now constituting me; or else 'the matter now constituting my body' is just another name for my body, this human organism that can survive the gain and loss of parts. On either alternative, to make something out of some matter is really to cause the matter to be replaced by something that is *not* constituted by *it*, since it is no more. On the face of it, neither alternative does justice to the obvious facts: that there is some matter constituting my body now and that this very matter does not have a human form at every time it exists. (2003: 517)

On what appears to be the most favorable interpretation of the view, the Thomist embraces the second horn of Zimmerman's dilemma: there is not some matter composing my body now which does not have a human form at every time at which it exists. When I breathe in an oxygen atom, that atom ceases to exist and is replaced by something else, albeit something that has the causal powers of the original now directed by the biological imperatives of the whole. When that new virtual oxygen atom ceases to be part of me, it too is replaced by something else, presumably a new oxygen atom. As Zimmerman indicates, this seems like a very implausible view. If someone ingests a radioactive dye, we can witness what appear to be the very same radioactive markers being caught up into his or her metabolic processes.

Thomists might reply that their view is not as implausible as it might at first appear since the original oxygen atom, the virtual atom which is a part of me, and the new oxygen atom that replaces the virtual one all have the same causal powers. As a result, there can be no observable differences among them, and hence their view accommodates all the empirical facts. In that case, though, Thomists appear to be introducing a difference without a difference, something that is a strike against theoretical

economy. Unless there is good reason to reject the alternative hylomorphic view, according to which one and the same oxygen atom persists both independent of me and as a part, it is difficult to appreciate why we should prefer the Thomistic view.

If there really were no way of accounting for the unity of a composite individual other than by endorsing the Thomistic view, then we would have good reason to reject the hylomorphic view I've defended. But is the Thomistic way the only way? Hitherto, I've characterized the unified character of diverse parts in terms of their contributions to the activity of a whole. The activity of a structured whole is not simple, but complex; that is why the whole requires parts: complex activities require complex composition. On the hylomorphic view I've defended, the complexity of activities is revealed through functional analysis. Mechanistic explanations of the activities of a complex whole reveal what Marmodoro calls the "roles" that are "allotted" to various parts (2013: 17). Things like oxygen atoms or electrons become parts by playing those roles. In doing so, they gain new properties they didn't previously have, properties that depend on the whole to which they belong, and that mark their contributions to its activities. Their powers are instead limited or directed by the wholes in which they are incorporated since those wholes have additional parts whose powers inhibit, enhance, or otherwise influence the powers of other parts. Individuals like electrons have natures that comprise various powers. Some of those powers they manifest only when incorporated into a structured whole, others they manifest only when they are not incorporated into structured wholes, and yet others they manifest in both cases. Because things like electrons have natures of this sort, composite wholes are able to harness the powers of those things in ways that contribute to their own biological ends. What Thomists have to provide is a clear argument for why an account of this sort isn't sufficient to account for the unity of a whole. Why suppose that a thing can play a part-defining role only if it is incapable of existing apart from a whole? It is easier to appreciate in the case of parts like the heart since they (or analogues of them) never exist apart from more inclusive wholes on the hylomorphic view. But in the case of, say, electrons, why should we suppose that they cannot play the roles dictated by the activity of the whole? Why suppose that they must cease to exist and be replaced by essentially biological proxies?

If we grant that the Thomistic view is sufficient to account for the unity of a whole, the question remains whether it is necessary. Establishing this would require an argument to the effect that views like the hylomorphic one I've defended aren't up to the task, and that is what we have yet to see Thomists provide.

6.6 Levels in the Natural World

One implication of the hylomorphic view of composition is that parts stand in hierarchical relationships. What qualifies an object, *a*, as a proper part of a more inclusive whole is that it contributes to the activity of that whole. If there are other objects, $b_1, b_2, \ldots, b_n$ that are proper parts of *a*, then $b_1, b_2, \ldots, b_n$ qualify as parts

because they contribute to *a*'s activity in turn, and the same is true mutatis mutandis if there are objects that compose the *b*s. In general, if $x_1, x_2, \ldots, x_n$ compose y, then $x_1, x_2, \ldots, x_n$ belong to a lower level than y. Hylomorphism is thus committed to a multilayered or multilevel worldview, a view according to which there are various layers or levels of reality (Jaworski 2011: 285–8). There is nevertheless an important difference between hylomorphism and other multilevel worldviews.

Many multilevel views define levels *globally*, relative to branches of science whose categories cut across the features that distinguish one kind of living thing from another. According to these views, levels correspond to different scientific disciplines (Oppenheim and Putnam 1958). The lowest level corresponds to fundamental physics, the next to atomic physics, then chemistry, biology, psychology, and finally, social science. Something's position in the hierarchy is determined, as it is on the hylomorphic view, by composition. Entities postulated by lower-level frameworks compose entities postulated by higher-level ones: fundamental physical particles compose protons and neutrons, protons and neutrons compose atoms, atoms compose molecules, molecules compose tissues, tissues compose organs, organs compose organisms, and organisms compose economic, political, and other social systems. Because levels correspond to scientific disciplines on these views, there is a single hierarchy of levels for the entire natural world irrespective of the specific kinds of organisms that exist. The biological, chemical, and physical levels we find in dogs, say, are the same as the biological, chemical, and physical levels we find in humans, or spiders, or cats. Hylomorphism does not deny that some levels cut across kind-specific boundaries. Fundamental physics, for instance, describes the basic materials that are liable to further structuring in entities of all kinds. But according to hylomorphists, not all levels are like this, not all can be defined globally; some can only be defined *locally*.

According to hylomorphists, different kinds of living things comprise different hierarchies of levels. The levels we find in dogs might be very different from those we find in cats. The reason is that levels are defined primarily by the relation of parthood or composition, and on the hylomorphic view, composition consists in the contributions that parts make to the activities of the wholes to which they belong. Consequently, if dogs and cats engage in different activities or if their parts contribute to their respective activities in different ways, then dogs and cats will comprise different kinds of parts, and that means they will comprise different hierarchies of levels. Moreover, it seems to be an empirical matter of fact that different kinds of living things engage in different activities and have parts that contribute to their respective activities in different ways. Consider an example.

Humans and spiders engage in different kinds of activities. Spiders but not humans spin webs, and conversely, humans but not spiders speak and sing. These differences in human and spider activity are reflected in differences in human and spider parts. Spiders but not humans have spinnerets, and humans but not spiders have vocal cords. When describing and explaining the behavior of humans and of

spiders, then, we use different predicates and terms to refer to or express their different activities and the different kinds of parts those activities involve. Moreover, even when we use the same predicates and terms to describe the activities of both, there are still significant differences in what these predicates and terms refer to or express. Consider eating. We can apply the term 'eating' both to an activity in which humans engage and to an activity in which spiders do. But while eating in humans and eating in spiders have some features in common, there are significant differences that become apparent as soon as we consider how members of each species eat. Humans eat by biting off morsels of food with their incisors and chewing them with their molars. Human saliva contains enzymes that partially digest the morsels which are eventually swallowed and carried to the stomach by the peristaltic motion of the esophagus, where they are further digested by acids and enzymes. Spiders, on the other hand, eat by regurgitating stomach fluids onto their prey while grinding at the prey with their chelicerae (their jaws). The chewing action of the chelicerae, in combination with the enzymes in the stomach fluid, break down the prey's tissues into a soupy mixture, which is sucked through the spider's mouth by the pumping action of its stomach while hairs around the mouth filter out any unliquefied pieces. Despite the common label, human eating and spider eating are very different activities that comprise very different subactivities and subsystems. Humans do not have chelicerae, nor do spiders have molars. The human stomach does not engage in pumping as the spider stomach does, nor does the spider bring food to its stomach by peristalsis as the human esophagus does. Humans and spiders, then, are composed of different kinds of parts suited to different kinds of activities.

But if levels are defined by composition, and humans and spiders are composed of different kinds of parts, then humans and spiders will comprise different hierarchies of levels. Each animal will comprise a hierarchy unique to members of its kind. Hylomorphism thus defines levels in a local, kind-specific way. Except for fundamental physics, and perhaps some other lower-level disciplines, the levels comprised by one kind of organism may not be found in organisms of any other kind. As a result, the hylomorphic multilevel worldview countenances a plurality of level hierarchies, each of which is tied to a distinctive kind of living thing. What emerges is a picture of organisms as multistructure complexes, each of which comprises a kind-specific hierarchy of structural levels (Figure 6.1).

Hylomorphism's local definition of levels converges with some recent work in the philosophy of science. Philosophers of neuroscience such as William Bechtel and Carl Craver have reached the conclusion that a local definition of levels is the most natural way to make sense of the explanations used in neuroscience. Craver, for instance, says the following:

> [L]evels of mechanisms are not monolithic divisions in the structure of the world . . . [They] are far more local than the monolithic image suggests. They are defined only within a given compositional hierarchy. Different levels of mechanisms are found in the spatial memory

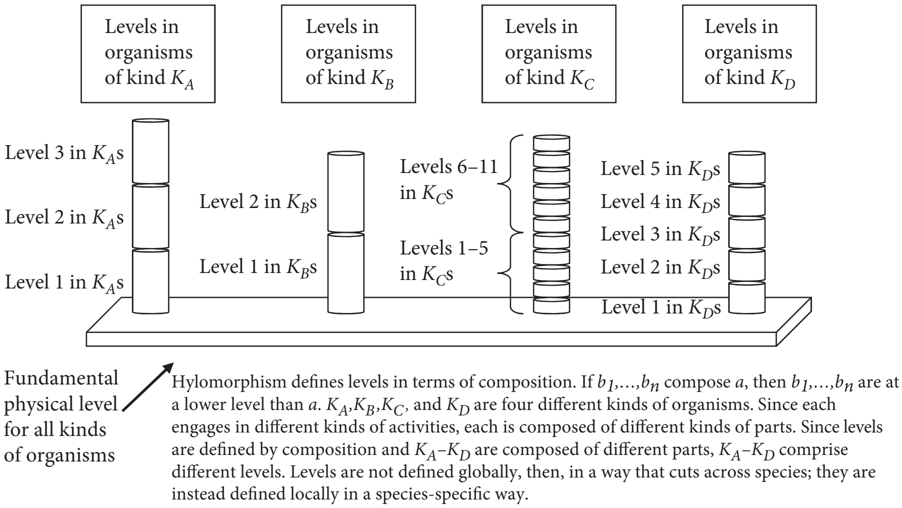


Figure 6.1 The hylomorphic multilevel worldview: locally defined levels.

system, the circulatory system, the osmoregulatory system, and the visual system. How many levels there are, and which levels are included, are questions to be answered on a case-by-case basis by discovering which components at which size scales are explanatorily relevant for a given phenomenon. They cannot be read off a menu of levels in advance. (2007: 190–1)

Consider now an objection to the hylomorphic view. John Heil (2003) has argued that multilevel worldviews are committed to the existence of higher-level properties, but, he says, the existence of higher-level properties is problematic. Among other things, it is unclear how such properties could be causally efficacious (see Section 5.3). As a result, multilevel worldviews—including ostensibly the hylomorphic worldview—implies that higher-level properties are epiphenomenal, an unacceptable result.

Heil's argument, however, does not apply to the hylomorphic worldview. The reason is in part terminological. We noted in Section 5.3 that calling a property 'higher order' is not the same as calling it 'higher level.' Levels can be introduced using any number of principles for hierarchical organization. To speak of higher and lower orders, on the other hand, is to speak of a specific kind of hierarchical principle, one involving logical operations. Higher-order properties are logical constructions that are defined using the operation of quantification; they are properties whose definitions quantify over other properties. If F_1, F_2, . . . , F_n are properties, we might define a property, *G*, by stipulating that something has *G* if and only if it has some *F*-property or other, or we might define *H* by stipulating that something has *H* if and only if it has some *F*-property that satisfies a certain condition *C*. Putnam (1970: 313–14) appears to have been the first to apply the term 'higher order' to logical constructions of this sort in the philosophy of mind (he took himself to be using the

term the way Russell and Whitehead did in *Principia*). Heil departs from this terminology. Instead of using the term 'higher order,' he uses the term 'higher level.' Consequently, when he speaks of views that endorse higher-level properties, he has in mind views that take the properties postulated by the special sciences to be logical constructions whose definitions quantify over lower-order properties. This is not the view of properties endorsed by hylomorphists.

We saw in Section 6.3 that according to hylomorphists, higher-level properties are first-order properties in their own right. It is this claim that distinguishes their view from functionalist views of all stripes. Functionalists claim that descriptions of higher-level phenomena are simply abstract descriptions of lower-level occurrences. Hylomorphists, by contrast, claim that higher-level descriptions correspond to distinctive natural structures that factor into descriptions and explanations of living behavior in ways that cannot be eliminated, reduced to, or paraphrased in favor of lower-level descriptions and explanations. We saw in Section 5.3, moreover, that hylomorphists reject the existence of higher-order properties in general. Hylomorphists do not deny that it is possible to define predicates and terms using quantification. What they deny is that the predicates and terms we define in this way must correspond to real properties, ones that make a causal difference to their bearers. This is simply an implication of the claim that properties are sparse, which was defended in Chapter 2. If properties are sparse, the only properties that exist are ones that enable individuals to enter into causal relations, and we cannot bring properties in this causal sense into existence merely by defining certain predicates and terms. For this reason, hylomorphists eschew talk of higher-order properties and speak instead of higher-order predicates, or terms, or descriptions (cf. Kim 1998: 104; Heil 2003: 45).

Philosophers who speak of higher-order properties reject a sparse conception of properties in favor of an abundant one. Examples include Prior et al. (1982), whose view was discussed in Section 5.3. According to them, "*every* expression like 'the property of being *F*' denotes a property" (1982: 254, emphasis in original). A conception of properties like this is what motivates the multilevel worldviews Heil finds troubling. Those views, he says, are motivated by what he calls the 'Picture Theory.' A central feature of the Picture Theory is a commitment to abundant properties:

> What exactly is the Picture Theory? . . . The core idea is that the character of reality can be 'read off' our linguistic representations of reality . . . A corollary . . . is the idea that to every meaningful predicate there corresponds a property. (2003: 6)

Hylomorphists agree with Heil's premise that higher-order properties are philosophically problematic. What they disagree with is the premise that any multilevel worldview must be committed to higher-order properties. They take their own view to provide a counterexample. Moreover, Heil (2003: 10, 73, 245) himself favors the idea that there are levels of organization or complexity, and suggests that, "The

truthmakers for claims about . . . people could turn out to be configurations of atoms in the void" (2003: 11). The question that arises from a hylomorphic perspective concerns how we are to understand the notion of configuration that Heil appeals to. If we take seriously the ontological implications of that notion, say hylomorphists, we end up endorsing something like the hylomorphic worldview.

6.7 Conclusion

The goal of this chapter was to situate the notion of structure introduced in Chapter 1 within the metaphysical framework developed in Chapters 2–5. The hylomorphic view defended here claims that structures are properties; they are powers of emergent individuals to configure the materials that compose them. Structured individuals engage in this configuring activity essentially and continuously. It is this activity that unifies them, and that enables them to persist through changes in their materials. This view of structure and composition is most easily understood as a way of developing van Inwagen's account of composition since lives on van Inwagen's view play many of the same roles that individual-making structures do: they are ontological and explanatory principles; they explain what things are and what things can do, and they account for the unity and persistence of those things.

There are nevertheless three differences between the hylomorphic view and van Inwagen's. First, the hylomorphic view is not committed to atomism as van Inwagen is; it remains neutral on the nature of fundamental physical materials. Second, hylomorphism is open to the possibility that there might be structured individuals other than living things. Third, hylomorphism endorses the existence of biofunctional parts such as eyes, hearts, and brains. This notion of parts gets its content in part from empirical disciplines that employ the method of functional analysis, a method that analyzes the activities of complex wholes into simpler subactivities performed by simpler subsystems. Parts on this view are hierarchically arranged systems and subsystems that are identified and individuated by the contributions they make to whole organisms. To use Furth's (1978: 638–9) analogy: if organisms are warps or vortices in the overall flow of matter and energy in the universe, then biofunctional parts are mini-warps or vortices—subpockets of order within the more inclusive wholes.

John Heil has argued against hierarchical views of nature. His argument, however, targets views that postulate higher-order properties, and the hylomorphic view does not postulate them. We saw in Section 5.3 that hylomorphists reject higher-order properties. Consequently, Heil's argument does not apply.

It is now time to defend the hylomorphic view of individual-making structure and composition against some objections. This is the goal of the next chapter.

7

The Problems of Composition

Chapter 6 was concerned with articulating the hylomorphic view of composition. This chapter is concerned with defending it. I will begin by considering arguments that have been advanced against the existence of parts like eyes, hearts, and brains—parts to which hylomorphists appear committed. Next, I will consider objections to van Inwagen's view of composition. Because the hylomorphic view is similar to van Inwagen's, it's important to consider whether these objections apply to the hylomorphic view as well. One objection concerns the possibility of atomless gunk, infinitely divisible stuff. A second concerns the Denial, the claim that there are no objects such as tables, chairs, and mountains. I will argue that hylomorphists can handle these objections at least as well as van Inwagen himself. Finally, I will consider a general challenge to views that endorse restricted composition: the vagueness argument. I will argue that hylomorphists have a plausible way of responding to it.

7.1 The Body-Minus Problem

There are at least two arguments that purport to show that there are no such parts as eyes, hearts, and kidneys. The first is advanced by Peter van Inwagen (1981). Consider again the many ways it is possible to divide human organisms into parts. According to van Inwagen, there is no principled reason for choosing to divide an organism into the kinds of functional parts hylomorphists recognize as opposed to, say, parts on a metric grid. Functional divisions do not reflect anything deep in reality, only "a linguistic accident" that aims to satisfy various interests we happen to have (1981: 127). In support of this claim, van Inwagen asks us to imagine a race of rational beings who raise humans as meat animals but who for religious reasons never eat human left legs. We can easily imagine these beings having a term that refers to left-leg-complements—that is, to the entirety of a human organism minus its left leg. This race of beings would have a set of descriptive interests different from ours, and would therefore divide humans into parts in a way different from the way we customarily do. Van Inwagen concludes that hands, eyes, hearts, and other parts postulated by science or commonsense are arbitrary parts: there is no principled reason for choosing to postulate them instead of parts of various other sorts. But arbitrary parts do not exist, says van Inwagen, for their existence would lead to a well-known problem with material constitution, the *body-minus problem* (Wiggins 1968; Burke 1994).

We can reconstruct the body-minus problem as follows. Suppose that (1) Descartes exists before *t*, and also (2) a proper part of Descartes, his left leg, exists before *t*. (3) If Descartes' left leg exists before *t*, then another part of him, Descartes-minus-his-left-leg (call it 'D-minus') exists before *t*. Imagine, however, that (4) Descartes survives the amputation of his left leg at *t*. Since, (5) D-minus also survives the amputation of Descartes' left leg at *t*, it seems that after *t* Descartes must be identical to D-minus. The reason is that (6) after *t* Descartes and D-minus have the same size, shape, position, orientation, attitude, mass, velocity, and color, but (7) two objects cannot have the same size, shape, position, orientation, attitude, mass, velocity, and color; hence, Descartes and D-minus must be identical.[1] The problem is that the identity of Descartes and D-minus appears to violate the axioms of identity. Because Descartes used to have two legs but D-minus didn't, for instance, it appears to violate Leibniz's law. Moreover, as van Inwagen argues, the identity of Descartes and D-minus also appears to violate the transitivity of identity. The foregoing premises imply all of the following claims:

(i) The thing that was D-minus before t = the thing that was D-minus after t;
(ii) The thing that was D-minus after t = the thing that was Descartes after t;
(iii) The thing that was Descartes after t = the thing that was Descartes before t;
(iv) The thing that was D-minus before $t \neq$ the thing that was Descartes before t.

If (i)–(iii) are true, the transitivity of identity implies that (iv) must be false.

Van Inwagen's solution to the body-minus problem is to deny that Descartes has any proper parts other than fundamental physical particles and individual cells. This implies that claim (2) is false. If there is no such thing as Descartes' left leg, then it no longer follows that there is such a thing as D-minus, and if there is no such thing as D-minus, the body-minus problem never gets off the ground. Since hylomorphists endorse biofunctional parts, this solution is not open to them; they must solve the body-minus problem a different way. They accept claim (2) and look instead to reject claim (3): they look to deny that the existence of biofunctional parts implies the existence of biofunctional part-complements. Left legs might exist, but it does not follow that left-leg-complements such as D-minus do.

Van Inwagen defends (3) on the grounds that postulating legs but not leg-complements would be arbitrary. After all, Descartes' left leg and D-minus comprise exactly the same kinds of physiological processes—cellular respiration, cellular replication, protein synthesis, and so forth. What principled basis could there be for claiming that the physiological processes in Descartes' left leg constitute a life while those in D-minus do not? The burden for any solution to the body-minus problem that rejects (3) is to provide a principled basis for distinguishing legs, hands, and

[1] Burke (1994) and Wiggins (1968) appeal to different principles here: Burke, to the principle that two things cannot share all the same parts at the same time, and Wiggins, to the principle that two things of the same sort cannot occupy the same place at the same time.

hearts from leg-, hand-, and heart-complements. The view of composition described in Sections 6.3–6.5 provides such a basis. According to that view, not all principles of part identity and individuation are created equal. Some ways of dividing organisms into parts are privileged, in particular, the ways implied by our best empirical descriptions, explanations, and methods. Privileging these descriptions, explanations, and methods is an implication of ontological naturalism (Section 1.2).

Ontological naturalism implies that if our best empirical descriptions, explanations, and methods divide organisms into biofunctional parts but not biofunctional part-complements, then we have good reason to think the former parts exist but not the latter. Do our best empirical descriptions, explanations, and methods divide organisms into biofunctional parts in fact? We saw in Chapter 6 that the answer appears to be yes: biologists, neuroscientists, psychologists, and others who are concerned with giving empirical descriptions and explanations of living behavior tend to view organisms as complex systems with parts that are individuated functionally in ways discovered through functional analysis. Philosophers of neuroscience such as William Bechtel and Carl Craver sometimes call complex systems of this sort 'mechanisms.' Something qualifies as a part of a mechanism, they say, only if it performs an operation that contributes to the activities of the whole:

> A mechanism consists of a particular set of parts that carry out specific operations, organized so as to produce a given phenomenon . . . Parts of mechanisms are not just any physically separable parts of a mechanism—rather, they are *working parts*, parts involved in the operations. (Bechtel 2008: 4, 14)

> A structure within the mechanism may be well delineated (it has boundaries, continues to exist over time, is differentiated from the things around it, etc.). However, if it does not perform an operation that contributes to the realization of the phenomenon, it is not a working part of that mechanism. For example, while the gyri and sulci of the brain are well delineated, they are not working parts of the brain but byproducts of the way brains fold to conserve the length of axons. (Bechtel 2007: 180)

> [W]hat is a mechanism? . . . [I]t is a set of entities and activities organized such that they exhibit the phenomenon to be explained . . . Entities are the components or parts in mechanisms . . . [T]he entities and activities in mechanisms are organized together spatially, temporally, causally, and hierarchically . . . Levels of mechanisms are levels of composition, but the composition relation is not, at base, spatial or material. In levels of mechanisms, relata are behaving mechanisms at higher levels and their components at lower levels . . . X's φ-ing is at a lower mechanistic level than S's ψ-ing if and only if X's φ-ing is a component in the mechanism for S's ψ-ing. (Craver 2007: 2–5, 188–9)

Observations like these about the kinds of parts postulated by actual work in the biological sciences give us good reason to think that our best empirical descriptions, explanations, and methods postulate biofunctional parts like eyes, legs, and hearts. They also give us good reason to think that those descriptions, explanations, and methods do not postulate eye-, leg-, and heart-complements like D-minus. The reason is that those descriptions, explanations, and methods individuate parts in

ways that are too fine-grained to accommodate parts like D-minus—parts that would include within their boundaries segments of many diverse subsystems: the circulatory system, the reproductive system, the visual system, and so on. It is possible that biologists, neuroscientists, and others might eventually discover that human behavior is described and explained more effectively using radically different principles of part identity and individuation, but in the absence of compelling reasons to think this will happen, our best current practices provide some reason to think that there are no such parts as D-minus.

The foregoing considerations provide hylomorphists with a way of solving the body-minus problem. If there are hands, legs, and hearts, but not hand-, leg-, and heart-complements, then claim (3) of the body-minus problem is false. The existence of biofunctional parts like Descartes' left leg does not imply that there are biofunctional part-complements like D-minus. The foregoing considerations also give us good reason to think that biofunctional part-complements like D-minus do not exist. As a result, the body-minus problem never gets off the ground: there can be no question of whether Descartes is identical to D-minus, for D-minus does not exist.

Given reasonable assumptions, this solution to the body-minus problem implies that (5) is false as well. Claim (5) says that D-minus survives the amputation of Descartes' left leg. If we assume that a statement of the form '*x* survives *y*' implies that *x* exists, then claim (5) is false if D-minus does not exist. Since the hylomorphic solution implies that D-minus does not exist, it also implies that (5) is false.

Consider now an objection to the hylomorphic solution. Hylomorphists have not succeeded in solving the body-minus problem, the objection says, for it is possible to reformulate the problem in a way that spells trouble for the hylomorphic view. Suppose that instead of D-minus, which admittedly stands little empirical chance of qualifying as a biofunctional part, we take something that uncontroversially qualifies: Descartes' brainstem, B. Why the brainstem? Eric Olson (1996) has argued that there is good empirical reason to suppose that the brainstem is the biofunctional part with which at the barest minimum a human animal can survive. Descartes survives if and only if his brainstem does. Let us call this the 'brainstem survival thesis.' Suppose that Olson is right about it. We can now reformulate the body-minus problem as follows:

(1′) Descartes exists before *t*, and (2′) a proper part of Descartes, B, exists before *t*. Due to some catastrophe, however, every proper part of Descartes except B and its proper parts is destroyed at time *t*. (3′) If B is the bare minimum biofunctional part with which a human animal can survive, then Descartes survives the catastrophe. So Descartes exists after *t*. But (4′) B also survives the catastrophe, so it too exists after *t*. Now, it seems that after *t* Descartes must be identical to B since (5′) after *t* Descartes and B have the same size, shape, position, orientation, attitude, mass, velocity, and color, but (6′) two objects cannot have the same size, shape, position, orientation, attitude, mass, velocity, and color. Hence, Descartes and B must be one, yet the identity of Descartes and B would violate Leibniz's law and the transitivity of identity for the reasons discussed earlier.

One advantage of van Inwagen's solution is that it elegantly solves both the earlier body-minus problem and this reformulated problem in the same way. In both cases, it denies that Descartes has a proper part, either his left leg or B. Hylomorphists, however, are not free to deny the existence of proper parts postulated by our best empirical descriptions and explanations of human behavior, and based on what's been said, that includes biofunctional parts such as brainstems. Hylomorphists must therefore solve the reformulated problem in a different way.

Hylomorphists are committed to the same constraints on identity and composition that van Inwagen is. They claim that identity is absolute, not relative, that there is only one thing located exactly where Descartes is, and that Descartes is wholly present at every time at which he exists. Consequently, hylomorphists cannot admit that there are two things located exactly where Descartes is, nor can they appeal to temporal parts, or to the idea that Descartes and B are identical relative to some sortals or times but not relative to others. It seems, therefore, that they must either deny that Descartes exists after *t*, or else deny that B does. Since by hypothesis a human animal survives the catastrophe, there is a strong presumption in favor of Descartes existing after *t*—what other human animal could it be? If hylomorphists stipulate that Descartes exists after *t*, then they should argue that B ceases to exist at *t*; that is, they should reject claim (4′).[2] Here is one way of articulating this kind of solution.

Suppose that 'B' rigidly designates Descartes' brainstem, a proper part of him which immediately prior to *t* is composed of objects $f_1, f_2, \ldots, f_n$—the cells, organelles, molecules, and so forth that would be revealed through functional analyses of B's activities before *t*. After time *t*, objects $f_1, f_2, \ldots, f_n$ no longer compose a proper part of Descartes, they compose only Descartes himself. After *t*, therefore, what had been a proper part of Descartes, his brainstem, no longer exists, even though Descartes does.

One worry about this solution is that it is incompatible with the brainstem survival thesis mentioned a moment ago, the thesis that Descartes survives if and only if his brainstem does. Since there are empirical reasons to think the brainstem survival thesis is true, hylomorphists should try to find a way of accommodating it or something very much like it. Here is one suggestion. Suppose that $f_1, f_2, \ldots, f_n$ are related to Descartes in the following way: so long as $f_1, f_2, \ldots, f_n$ continue to contribute to Descartes' overall operation as they have done hitherto, Descartes will continue to exist, but if any of them is damaged or destroyed at a time, and is

[2] Howard Robinson has suggested to me in conversation that hylomorphists might do better by rejecting (6′), the claim that two objects cannot have the same size, shape, position, orientation, attitude, mass, velocity, and color. This would involve rejecting the constraint that there is only one thing located where Descartes is. The worry I have about this strategy is that if we employ it in the brainstem case, there seems nothing to prevent someone from employing it in other cases. Suppose, for instance, that we take 'Descartes' body' to refer to the whole collection of Descartes' parts prior to *t*, and that we claim that these parts compose Descartes; we claim, in other words, that Descartes is composed by his body. The result seems little different from the collocation views favored by Lynne Rudder Baker (2000), Kit Fine (1999; 2008), Mark Johnston (2006), and E. J. Lowe (1996).

thus incapable of performing its contributing subactivity, then Descartes will cease to exist at that time. Collectively, then, $f_1, f_2, \ldots, f_n$ constitute a core of functional components that Descartes needs to exist, and that are sufficient to enable him to exist. They are in this sense the barest minimum functional parts with which Descartes can survive. Since objects $f_1, f_2, \ldots, f_n$ composed Descartes' brainstem prior to t, someone might be tempted to say that Descartes' brainstem is the bare minimum that is needed for Descartes to survive. Strictly speaking, however, this is false. What is true instead is that Descartes survives if and only if $f_1, f_2, \ldots, f_n$ do, and $f_1, f_2, \ldots, f_n$ needn't compose a brainstem. After the catastrophe, Descartes survives along with $f_1, f_2, \ldots, f_n$, but his brainstem does not. Moreover, because Descartes has no proper parts other than those which formerly composed his brainstem, his own activities are now solely the activities to which $f_1, f_2, \ldots, f_n$ contribute, the activities previously performed by his brainstem. Because the range of Descartes' activities has been diminished in this way, someone might be tempted to say that Descartes has become his brainstem. We have seen, however, that this is a mistaken way of describing the situation, one that generates the body-minus problem. The correct way of describing the situation is instead to say that Descartes survives the catastrophe, along with the proper parts that used to compose his brainstem, but that now compose only him.

Critics might complain that this solution is implausible; in particular, it is implausible to suppose that Descartes' brainstem would disappear even though all the parts that previously composed it remain where they are. There are at least two things hylomorphists can say in response. First, it is important to emphasize that what critics must find objectionable here is not that Descartes' proper parts should cease to exist in the process of whittling him down. Any whittling process is likely to result in a thing losing parts. What they must find objectionable is rather that in the brainstem case the parts that previously composed Descartes' brainstem retain all the intrinsic properties and spatial relations to each other that they previously had. To the extent that the objection is motivated by this, it appears to be tacitly committed to the idea that the intrinsic properties and the spatial arrangement of $f_1, f_2, \ldots, f_n$ are sufficient for those things to compose a brainstem. But the biofunctional account of composition suggests otherwise. For on that account, $f_1, f_2, \ldots, f_n$ compose something only if they contribute to the activities of the whole they compose. Suppose, then, that brainstems are essentially things that coordinate the activities of diverse organ systems. Van Inwagen himself appears to endorse a principle like this:

> Consider . . . a freshly severed head [and] a newly headless body . . . [S]uppose that an object of each of these types has been provided with an appropriate "life-support system." In my view the severed head is a genuine living organism and the headless body is not . . . In virtue of what does the activity of the simples composing the one constitute a life and the activity of the simples (virtually) composing the other fail to constitute a life? . . . [T]he difference is this: Give the severed head the proper environment and it will maintain itself, but the headless body will need a constant supply of "instructions" in the form of electrically transmitted information. Unlike the head, it will not be able to coordinate its activities. (1990: 177–8)

The crucial difference between the head and the rest of the body, on van Inwagen's view, has to do with the head's ability to coordinate the activities of diverse parts. Likewise, according to van Inwagen, if the activities of the *x*s constitute a life, then the *x*s cannot be causally isolated, they must interact (1990: 150). This suggests that being a part consists partly in standing in certain relations to other parts.[3] Suppose that this is true, and that being a brainstem consists partly in coordinating the activities of diverse organ systems. When Descartes is whittled down to $f_1, f_2, \ldots, f_n$, he no longer has any diverse organ systems to coordinate. In that case, however, it is not implausible to claim that he does not have a brainstem among his proper parts.

Second, hylomorphists can argue that even if it is implausible to suppose that Descartes' brainstem ceases to exist when he is whittled down, it is nevertheless no more implausible than the alternatives. It is no more implausible, for instance, than claiming that Descartes' brainstem never existed at all. But if this solution to the body-minus problem is no more implausible than the alternatives, then the charge of implausibility loses its force.

The foregoing considerations show that hylomorphists have a way of dealing with one of van Inwagen's arguments against the existence of eyes, hearts, and similar biofunctional parts. But van Inwagen suggests a second argument as well. Lives are jealous, he says—so jealous that it is impossible for two lives to overlap unless one of them is subordinate to the other. The only case in which this can occur, he says, is a case in which the life of an individual cell is subordinate to the life of a multicellular organism (1990: 89). Since eyes, hearts, and similar biofunctional parts are not individual cells, they cannot be subordinate to the life of a multicellular organism. But neither are they multicellular organisms; if they exist at all, they must be proper parts of multicellular organisms. Consequently, eyes, hearts, and similar biofunctional parts must not exist.

Hylomorphists can respond to this argument by challenging the premise that the only cases in which lives are subordinate are cases involving individual cells. Clearly, van Inwagen thinks that if there were subordinate lives other than those of individual cells, it would undermine the jealousy of lives. He does not argue for this claim, however, nor does he explain why the jealousy of lives is not undermined by the subordinate lives of individual cells. It seems plausible to think that his reasons might concern the unifying roles that lives are supposed to play. Recall that *lives count* on van Inwagen's view: a life is supposed to explain why diverse things compose a single unified whole (Section 6.2). The larger the number and variety of things involved, the more unifying work the notion of a life must be pressed into service to perform. Someone might feel that lives could plausibly be asked to unify diverse cells and physical particles, but that it would take things too far to ask them to unify

[3] We find a similar idea in Aristotle's *Parts of the Animals* 654a32ff. The idea that a part might be defined in part by its relations to other parts is the reason he claims that strictly there are no individual bones, but only a skeletal system.

multicellular tissues, organs, and organ systems in addition. Since individual-making structures are supposed to count on the hylomorphic view, the same point, it seems, would apply to it.

One way hylomorphists can respond is by elaborating the notion of subordination to include a notion of subordinate unity. A multicellular organ enjoys a unity analogous to that of a multicellular organism. It is unified in a way that enables it to perform the subactivity that qualifies it as a proper part of a whole. Hylomorphists might argue, moreover, that this subordinate unity explains why so-called "severed parts" sometimes display behaviors similar to the behaviors of parts in situ. A heart removed from the chest cavity continues to beat (or more precisely, the materials that used to compose a heart continue to change position relative to each other in a manner that closely resembles the beating of a heart). This lifelike behavior can be explained as a vestige of the unity those materials once possessed when they composed a heart. That vestige of unity can be used to explain in turn why it remains possible for those materials to recompose a heart when they are transplanted to a new individual even if the same materials are not transplantable in the same way if smashed or incinerated. The latter processes disrupt the vestige of unity those materials enjoyed as parts, but if that vestige remains intact, then it remains possible for the materials to take up once again the kind of role they formerly played. Suppose, then, that parts enjoy this kind of subordinate unity. Because the unity is subordinate to the unity of the wholes to which they belong, their existence should not compromise the jealousy of lives any more than the existence of individual cells does. Consequently, if van Inwagen is willing to countenance the existence of individual cells, nothing stands in the way of hylomorphists countenancing the existence of legs, hands, hearts, and other biofunctional parts.

One final clarification before continuing: The arguments above focus on parts such as legs, hands, and hearts. But someone might accept that there are biofunctional parts while yet denying that legs, hands, and hearts are included among them. Perhaps many of the parts we recognize in our pedestrian dealings will not appear on the official part list that is drafted in light of our best empirical descriptions and explanations of human behavior. This is a point hylomorphists readily concede. The important takeaway from the foregoing discussion is that biofunctional parts are not limited to fundamental physical materials and individual cells, as van Inwagen would have it. There are other kinds of biofunctional parts in addition. Which kinds of additional biofunctional parts there are is something that on the hylomorphic view is to be decided empirically.

7.2 The Problem of Too Many Thinkers

Another kind of argument against biofunctional parts is the problem of too many thinkers. It has been advanced by Trenton Merricks (2001a: 49–53), and independently by Eric Olson (1995: 187). Here is one version of it. The hylomorphic account

of composition implies both that there are organisms, such as you and I, and that there are biofunctional parts of those organisms, such as eyes, hearts, and brains. Suppose for the sake of argument that my brain and I both exist. (1) If my brain and I both exist, then both of us are rational, conscious beings, and (2) if both of us are rational, conscious beings, then assuming that I am not identical to my brain, it follows that located in the region of space I occupy there are two rational, conscious beings. But (3) this implication is absurd. Among other things, it appears to imply that I cannot know whether I am William Jaworski or merely William Jaworski's brain. It seems reasonable to assume that if my brain can think, then its thoughts are no different from my thoughts. If I want *x*, then it wants *x*; if I hope that *p*, then it hopes that *p*. Consider, then, my belief that I am William Jaworski. It seems that my brain must also have this belief. The only difference between me and my brain is that my belief is true, whereas its belief is false. This implication by itself will strike many people as absurd, but consider another. Given that my brain and I have all the same thoughts, how do I really know that I am William Jaworski and not merely William Jaworski's brain? Each of us believes that he is William Jaworski, but how can either of us know that he really is William Jaworski? It seems that there is no way either of us can tell. Since the implications of this line of reasoning are absurd, we should reject the original supposition: we should deny that my brain and I both exist. But it is much more plausible to deny that my brain exists than it is to deny that I exist. Therefore, we must conclude that my brain does not exist. What is true of brains, moreover, is true of biofunctional parts in general. We thus have good reason to conclude that there are no such parts as eyes, hearts, and brains.

The crucial premise here is (1). It claims that my status as a rational, conscious being is yoked to my brain's status as a rational, conscious being. There are several ways of defending this premise. One is to claim that any features that are relevant to being a rational, conscious being are features that I would have to share with my brain if it existed. It seems, for instance, that my status as a rational, conscious being depends on me having proper parts spatially arranged brain-wise. From this one might generalize and claim that anything with proper parts spatially arranged brain-wise must be a rational, conscious being (Merricks 2001a: 49). Since my brain and I both have proper parts spatially arranged brain-wise, it follows that both of us must be rational, conscious beings. Why accept the general thesis that anything with parts spatially arranged brain-wise must be a rational, conscious being? Merricks suggests three arguments.

First, Merricks says that denying that brains are rational, conscious beings would yield unwarranted skepticism about who or what is conscious (2001a: 50). Why should we accept this? Merricks does not make the steps of his argument explicit, but perhaps he has in mind something like the following: (i) If not everything that has proper parts spatially arranged brain-wise is a rational, conscious being, then knowing that something has proper parts spatially arranged brain-wise does not provide sufficient grounds for knowing that it is a rational, conscious being. But (ii) if

knowing that something has proper parts spatially arranged brain-wise does not provide sufficient grounds for knowing that it is a rational, conscious being, then I can entertain serious doubts about whether, say, you are a rational, conscious being. But surely (iii) this kind of skepticism is unwarranted. Hence, knowing that something has proper parts spatially arranged brain-wise must provide sufficient grounds for knowing that it is a rational, conscious being, and that means in turn that everything that has proper parts spatially arranged brain-wise must be a rational, conscious being.

In response, hylomorphists can target Premise (ii). There are well-rehearsed considerations in the philosophy of mind that suggest that (ii) is false: knowing that someone or something is rational and conscious does not depend on knowing anything about its brain states; in fact, it does not depend on knowing that it has proper parts spatially arranged brain-wise (as opposed to otherwise) at all.

First, people such as Aristotle have believed that brains play no role in rational, conscious behavior. If belief is necessary for knowledge, then clearly these people did not know that having proper parts spatially arranged brain-wise was sufficient for something to be rational and conscious, yet they presumably knew just as much as we do about people being rational, conscious beings. If Premise (ii) is true, this knowledge becomes utterly inexplicable.

Second, the idea that it is possible to know that people are rational, conscious beings without knowing that they have proper parts spatially arranged brain-wise has been commonplace in the philosophy of mind since the advent of multiple realizability arguments in the late 1960s (Putnam 1967; Fodor 1968a, 1974; Block and Fodor 1972). According to these arguments, we can and typically do make accurate judgments about whether someone or something is rational or conscious without knowing any details about the internal physiological, mechanical, or other states that correlate sensory inputs to it with behavioral outputs from it. It is possible, for instance, that even though you are fully convinced that I am rational and conscious, you might still discover that I do not have a human brain in my skull, but a functionally equivalent Martian organ. If cases like this are possible, then knowing about rationality and consciousness is independent of knowing about the physiological underpinnings of rationality and consciousness. But if knowledge of the one thing is independent of knowledge of the other, if it is possible to know that S is rational and conscious without knowing anything about S's physiological make-up, then knowing that S is rational and conscious does not depend on knowing that S has internal parts spatially arranged brain-wise, nor does it depend on knowing that having proper parts spatially arranged brain-wise is sufficient for rational, conscious behavior. And if these things are epistemically independent of each other, then it is possible that I might have complete certainty about one of them while yet being completely ignorant of the other. It is possible, for instance, that I might be incapable of entertaining serious doubts about whether you are a rational, conscious being, and yet be completely ignorant that having proper parts spatially arranged brain-wise is

sufficient for knowing that something is rational and conscious. The foregoing considerations suggest that Premise (ii) of Merricks' first argument is false.

Merricks' second argument claims that any view denying the sufficiency of having proper parts spatially arranged brain-wise for being rational and conscious must explain why having proper parts spatially arranged brain-wise makes some things, such as you and I, rational and conscious, but not others, such as our brains (2001a: 51). Again, Merricks does not elaborate on this point, but presumably the worry is that a view that doesn't explain this would be unprincipled or ad hoc. This argument is closely related to Merricks' third. Any view denying the sufficiency of having proper parts spatially arranged brain-wise for being rational and conscious must explain what would happen if I were to survive being whittled down to just my brain. Yet, says Merricks, any attempt to explain this is bound to be problematic (2001a: 52–3). Merricks does not make the steps of this argument explicit, but perhaps the following is a fair reconstruction: The whittling case appears to be one in which a person becomes a brain, a thing composed entirely of proper parts spatially arranged brain-wise. Suppose, then, that (i) persons such as you and I are rational, conscious beings, but that (ii) brains are not rational, conscious beings. (iii) After being whittled down, you and I are identical to our brains. From these premises, it follows that either you and I cease to be rational, conscious beings, or else brains can be rational, conscious beings after all—contrary to claim (ii). But (iv) it is implausible to suppose that you and I would cease to be rational, conscious beings if whittled down. Hence, the argument concludes, brains must be rational, conscious beings if persons such as you and I are. Claims (i)–(iv) are all plausible, so any view that denies the conclusion must explain which of them is false, yet, says Merricks, denying any one of them is bound to be problematic.

We saw in Section 7.1 how hylomorphists approach whittling cases. Their approach implies that Premise (iii) is false. If I were whittled down, I would not become my brain as (iii) says; rather, say hylomorphists, my brain would cease to exist at some point in the whittling process. I, on the other hand, would survive and be composed only of those proper parts which had at one time composed my brain, but which no longer do. Merricks does not say why a response along these lines must be problematic. In the absence of a further argument, I conclude on the basis of the considerations advanced in Section 7.1 that this response is not as problematic as Merricks insists.

That same response can be used to deal with Merricks' second argument. Recall that according to hylomorphists, the only things that have brains are organisms. There can be no severed brains—no materials that are structured brain-wise—without there being an organism that configures them (Section 6.4). Moreover, we've seen that an organism that was whittled down to materials spatially arranged brain-wise would not become a brain on the hylomorphic view; it would remain a whole organism, albeit one that was shrunken and debilitated (Section 7.1). The materials such an organism would configure would no longer compose a brain, but

only the organism itself. Merricks' second argument requests an explanation for why having proper parts spatially arranged brain-wise should make some things rational and conscious but not others. The explanandum involves two kinds of cases: (a) cases in which there is a rational, conscious organism that has proper parts spatially arranged brain-wise, and (b) cases in which something is not a rational, conscious organism, but it nevertheless has proper parts spatially arranged brain-wise. Merricks challenges any view denying that having proper parts spatially arranged brain-wise is sufficient for being rational and conscious to explain why there are cases of sort (b), given that there are cases of sort (a).

Hylomorphists respond that when it comes to something being rational and conscious, what matters is not having parts spatially arranged brain-wise, but having the right kind of structure. Structure involves more than having a mere spatial arrangement of parts (Section 6.1, 6.3–6.4). The structures of living things are dynamic: each has a temporal dimension which comprises complex relations among its parts and programmatic sequences of changes over time. Consequently, something's having a brain-wise structure is different from some materials simply being spatially arranged brain-wise. Cases of sort (a) involve rational, conscious organisms whose individual-making structures confer on them their distinctive powers. Because they have those structures, they have various biofunctional parts, including, we might suppose, brains which embody their powers for rationality and consciousness. Consequently, nothing other than organisms with the right kinds of structures can have brains. There can be no standalone brains, but only brains that are proper parts of organisms.

Given the foregoing considerations, it is difficult to think of cases of sort (b). Since only structured individuals have proper parts on the hylomorphic view, cases of this sort would have to involve structured individuals (presumably organisms) that had proper parts spatially arranged brain-wise, but that lacked any rational, conscious capacities. Only two possibilities come to mind. The first would be cases in which people suffered neural trauma severe enough to rob them of their capacities for rationality and consciousness. The skulls of such people would arguably still contain proper parts spatially arranged brain-wise, but in the envisioned cases, the trauma would have destroyed the causal connections among those parts, and thereby made it impossible to impose a rational, conscious order on the manifestation of their powers. The clearest way for a hylomorphist to respond to these kinds of cases is to deny that they actually involve proper parts spatially arranged brain-wise. A brain-wise spatial arrangement of parts requires that the connections among various parts of the brain remain intact. In the case of a severe neural trauma victim, those connections have been destroyed. The mass of tissue housed in the skull is no longer spatially arranged brain-wise. It is spatially arranged in a way that superficially resembles a brain-wise spatial arrangement, but it is not a brain-wise spatial arrangement in fact.

The second case that comes to mind is a case in which we remove someone's brain. The resulting three-pound, pinkish-gray mass of folded tissue would undoubtedly be composed of materials spatially arranged brain-wise, but those materials, now

removed from their connection with the rest of the person, would not compose a rational, conscious organism. For materials to compose something on the hylomorphic view is for them to contribute biofunctionally to its activities. Insofar as the pinkish-gray materials removed from someone's skull would not be contributing to something's activities in this way, those materials would not compose anything. Since they would not compose anything, they would not be proper parts of anything. But cases of sort (b) are supposed to involve things that have proper parts spatially arranged brain-wise. As hylomorphists understand these cases, there would be nothing that had proper parts, and hence nothing that had proper parts spatially arranged brain-wise. The point is analogous to the earlier one about corpses (Section 6.4): a corpse is not an individual in its own right. 'Corpse' is a label for physical materials that used to be structured human-wise, but that no longer are. Likewise, the pinkish-gray materials removed from someone's skull would not compose anything: they would be materials that used to be structured brain-wise, that used to compose a brain, but that no longer do.

If hylomorphists are right, it is difficult to think of any case of sort (b). There are no things other than organisms that have proper parts spatially arranged brain-wise. But if there are no such things, then requesting an explanation for why there are such things is misguided. It is legitimate to request an explanation for why it is the case that *p* only if it is in fact the case that *p*. Since hylomorphists claim that nothing other than organisms have proper parts spatially arranged brain-wise, it is illegitimate to request of them an explanation for why having proper parts spatially arranged brain-wise fails to confer rationality and consciousness on something other than an organism—something such as a standalone brain. On hylomorphists' own terms, therefore, Merricks' request for an explanation is illegitimate—as illegitimate as requesting that a meteorologist explain how Zeus produces rain. Since Zeus does not in fact produce rain, the request is not legitimate at all. To insist that it is legitimate is to presuppose something that hylomorphists reject, and in that case, failing to satisfy the request cannot be used as a nonquestion-begging argument against the hylomorphic view.

Hylomorphists thus have principled reasons for rejecting the claim that having proper parts spatially arranged brain-wise is sufficient for something to be a rational, conscious being. In that case, however, nothing prevents them from rejecting Premise (1) of Merricks' argument, the claim that if my brain and I both exist, then both of us must be rational, conscious beings. Rather, according to them, rationality and consciousness are characteristics of whole organisms, not their proper parts. My brain and I both exist, and I am a rational, conscious being, but my brain is not.

7.3 Atomless Gunk

We saw in Chapter 6 that the hylomorphic view of composition is very similar to van Inwagen's. It is thus important to consider whether the objections that have been

advanced against van Inwagen's view can be advanced against the hylomorphic view as well. Van Inwagen's view of composition faces two well-known objections. The first concerns the possibility of atomless gunk—infinitely divisible material. Something composed of gunk would have no mereological simples as parts. It would be decomposable into parts, and those parts into further parts, and those parts into yet further parts, and so on ad infinitum. Since every part of an object composed of gunk would have proper parts, there would be no ground floor of proper parts, no foundational level consisting of mereological simples. It seems possible that there could be a world in which there were no living things or mereological simples, but only gunk. Ted Sider (1993) calls worlds like this, devoid of living things and simples, 'gunk worlds.'

Sider argues that the possibility of gunk worlds is problematic for van Inwagen for at least three reasons. First, van Inwagen's view has the implausible implication that in gunk worlds there are no material beings. Gunk worlds have neither living things nor mereological simples. Since these are the only material beings in van Inwagen's ontology, his view implies that a gunk world is completely devoid of material beings. Even a gunk world that has gunk spatially arranged in ways indistinguishable from what we would call 'chairs,' 'tables,' 'mountains,' and 'buildings' has no material beings. To make matters worse, van Inwagen cannot avail himself of the paraphrastic strategy described in Section 6.2. Talk of the chair in the corner cannot be paraphrased in a gunk world as talk of mereological simples spatially arranged in the corner chair-wise, for in a gunk world there are no mereological simples. Second, the actual existence of gunk is an empirical hypothesis. For all we know, the fundamental physical materials that actually exist are gunky instead of atomic. Determining whether this is the case is an empirical matter. To the extent that van Inwagen's account takes the issue to be settled already, it seems guilty of poaching on the reserves of science. Finally, van Inwagen's account is supposed to provide an answer to the SCQ. It seems plausible to demand, however, that an answer to the SCQ be true in every possible world. But the possibility of gunk worlds suggests that it is not necessarily the case that something is composed of things exactly if the activity of those things constitutes a life. Nor can van Inwagen opt for a more limited account that purports to answer the SCQ only for nongunk worlds. The reason is that van Inwagen argues that his account is the only one that provides a satisfactory answer to the SCQ, and his arguments against alternative answers do not depend on any assumptions about gunk. Imagine, then, an alternative account of composition that was true for gunk worlds. It would specify a condition, *C*, such that something would be composed by a quantity of gunk exactly if that quantity satisfied *C*. *C* would provide an account according to which some material beings had proper parts but were neither living things nor mereological simples. But if *C* were sufficient to provide an account of composition in gunk worlds, then it is plausible to suppose it would also be sufficient to provide an account of composition in nongunk worlds. In that case, however, van Inwagen's arguments would have to be flawed.

Dean Zimmerman (2003) has advanced another gunk-based objection to van Inwagen's view. It appeals to the possibility of *malleable gunk*, material that is divisible in an indefinite, perhaps an infinite, variety of ways. If there were such material, Zimmerman argues, then there would be no principled way of saying what quantities or what arrangements of gunky material composed something at a time. If I were composed of a quantity of that malleable gunky material, he argues, then the space I occupy could be divvied up in an indefinite variety of ways. The gunk composing me could be spatially arranged into tiny balls, tiny pyramids, tiny cubes—any way one likes. Suppose, then, that 'the matter composing me now' referred to certain quantities or arrangements of fundamental physical materials. Which quantities or arrangements? There would seem to be an indefinite, perhaps even an infinite, number of candidates. It could refer to the parcels of gunk spatially arranged ball-wise, or the parcels of gunk spatially arranged pyramid-wise, or one of any number of parcels of gunk spatially arranged in one of any number of different ways. Because there wouldn't be a principled way of choosing one referent for 'the matter composing me now' as opposed to another, the choice would be purely arbitrary—a very awkward result. In addition, Zimmerman argues that even if there were a principled way of choosing a referent for 'the matter composing me now,' the result would be a very implausible semantics for matter talk.

Can these objections to van Inwagen's view be made to apply to the hylomorphic view as well? It should be evident that the original gunk objection cannot. The reason is that unlike van Inwagen's view, hylomorphism is not committed to atomism. It is compatible with the possibility that fundamental physical materials should turn out to be gunky. As a result, the hylomorphic view does not imply that there are no material beings in a gunk world, for although there are no living beings in a gunk world, there are still fundamental physical materials. Moreover, hylomorphists are free to use a paraphrastic strategy like van Inwagen's. They can countenance talk of the chair in the corner—whether in a gunk world or not—as talk of fundamental physical materials (whatever those turn out to be) that are spatially arranged in the corner chair-wise. In addition, because the hylomorphic view remains neutral about the atomic or gunky nature of fundamental physical materials, it avoids the worry that it might end up poaching on the reserves of science. Finally, there is nothing to suggest that the hylomorphic view cannot be true in every possible world. It thus avoids the worries about infinitely divisible gunk that face van Inwagen's view.

What about the problem of malleable gunk? It claims that if there is malleable gunk, there must be multiple candidates for what 'the matter composing me at *t*' refers to, and no principled way of determining which among those candidates the term refers to in fact. Hylomorphists respond that if there is gunk, it cannot be malleable in the way the objection assumes. According to hylomorphists, I have no parts other than biofunctional ones: nothing composes me unless it contributes biofunctionally to my activities. Consequently, if we assume that a quantity of

gunky material is a part of me, then that material must compose a biofunctional part. It is possible for us to conceive of a variety of principles for part identity and individuation (Section 6.4). We can conceive of dividing human organisms not along biofunctional lines, but along purely spatial lines into halves or thirds, or the spatial parts yielded by a three-dimensional metric grid. But according to hylomorphists, principles of part identity and individuation like these fail to carve organisms at their joints: they fail to pick out the kinds of parts that really exist. The principles that succeed in doing so are the ones implied by our best empirical descriptions and explanations of living behavior, and there is good reason to think that those descriptions and explanations divide organisms into biofunctional parts (Sections 6.3, 6.4, 7.1). The malleable gunk objection assumes that there is no privileged principle of part identity and individuation along these lines, that all such principles could carve composite individuals at their joints equally well. Since the objection is based on an assumption that hylomorphists reject, it begs the question against them. It does not prove that their view is false; it rather assumes its falsity from the outset.

The hylomorphic view is immune to the foregoing gunk-based objections. Consider now a third gunk-based objection. Suppose that the hylomorphic view of composition is in fact correct, and that the only parts are biofunctional ones. Suppose, moreover, that fundamental physical materials are gunky in world *w*. These claims seem to imply that in world *w* biofunctional parts are themselves gunky, that the list of systems, subsystems, and sub-... sub-subsystems ends up being infinitely long, and that functional analyses of biological activities end up going infinitely deep. But this view of biofunctional parts seems highly implausible, says the objection. We typically do not suppose that functional differentiation goes infinitely deep, that there is or could be an infinite number of jobs for systems, subsystems, and sub-... sub-subsystems to perform. One reason we do not suppose this is that the functions performed by subsystems typically get simpler as functional analyses proceed. The operation of the circulatory system is simpler than the operation of the animal as a whole, the operation of the heart is simpler than that of the circulatory system, the operation of an individual muscle fiber is simpler than the operation of the heart, and so on. The increasing simplicity of lower-level biofunctional roles makes it seem implausible that functional analyses could continue indefinitely; it suggests instead that living things have a foundational level of biofunctional parts. Whether or not they do in a given possible world is an empirical matter, but a theory of composition shouldn't be incompatible with either empirical outcome; otherwise, it can be charged with poaching on the reserves of science—a worry about the hylomorphic account of living things analogous to the worry about van Inwagen's account of fundamental physical materials.

This third objection is based on the premise that functional differentiation does not go infinitely deep. Consequently, one way for hylomorphists to respond is to accept gunky biofunctional parts while yet denying that functional differentiation

goes infinitely deep.[4] To see how they might articulate a response like this, consider an analogy with artifacts.

Imagine that artifacts have functions, and that the function of a two-by-four wall stud is to support the weight above it. The upper and lower halves of the stud have functions of the same type, and the same is true of the halves of those halves, and the halves of those halves in turn, and so on. The stud is thus composed of functional parts and subparts. If fundamental physical materials are gunky, then those parts are divisible into further parts ad infinitum. Each part and subpart makes a different token contribution to the activity of the whole since each performs its function at a specific location that differs from the locations of the others. These different token contributions individuate the stud's parts. And yet, those token contributions do not differ from each other in type. Yet, each of the stud's parts and subparts perform the same type of function: each supports the weight above it. Because all of the stud's parts and subparts make the same type of contribution to the whole, functional differentiation does not continue ad infinitum, even if divisibility into parts and subparts does.

Suppose, then, that something analogous is true of organisms and their parts. Suppose that, say, an electron is a biofunctional part of me and that the electron is itself divisible into further parts and subparts ad infinitum. Suppose, moreover, that the parts and subparts of the electron all contribute to its negative charge in the same type of way—a simple way that is not analyzable into different types of functional contributions. In that case, there would be a ground level of biofunctional differentiation even though fundamental physical materials were gunky. Biofunctional parts would be infinitely divisible, but the types of ways in which those parts contributed to wholes would be finite; at a certain level of analysis, it would be possible for each biofunctional part to contribute to the activity of a whole in the same type of way as other biofunctional parts at that level and lower levels. The parts would be distinguished from each other because they would make different token contributions to the whole.

What would expressions like 'the matter composing me at *t*' refer to on this account? Hylomorphists could say that such expressions refer to things that are biofunctional parts of me but that can exist without being parts of me or anything else. Recall that on the hylomorphic view some things that are biofunctional parts cannot exist without belonging to more inclusive wholes (Section 6.4).[5] Hearts, hands, and eyes are examples. A heart, when removed from an organism, ceases to be a heart. The heart-shaped materials that remain no longer compose anything since

[4] Heil (2003: 174) suggests a similar point. He says that even if things are spatially divisible ad infinitum, this does not show that there are no mereological simples. This would follow if some parts are not individuated along purely spatial lines. In particular, he distinguishes purely spatial parts from what he calls 'substantial parts' (100), which are roughly individuated by the functions they perform.

[5] Fine (1999) refers to these as 'timeless parts.'

they are no longer being configured by an organism, and on the hylomorphic view, composition happens only when structured individuals configure materials. But according to hylomorphists, not all things are like the heart. The electron that depolarizes one of my neural membranes at t is a part of me at t, but unlike the heart, it can exist without being part of me or any other living thing: its status as a part is merely contingent. Hylomorphists can say that 'the matter composing me at t' refers to things like the electron—things that are parts of me but that need not be part of me or any other structured whole. For these kinds of entities, being a part is a contingent property (Section 6.4–6.5).

A semantics along these lines captures the intuition that expressions like 'the matter composing me at t' refer to things that are parts of me only contingently, an intuition expressed in statements like the following:

(1) The matter composing me now is not the same as the matter that composed me last year (because I am an organism, and organisms are constantly exchanging matter with their environments).
(2) The matter that composed me on July 3, 1997 is the same as the matter that is now composing that small tree (for after that matter ceased to compose me, it managed by a series of historical accidents to become integrated into the tree).
(3) The matter that composed me on July 3, 1997 is now dispersed throughout the biosphere.

On the proposed account, statement (1) says that the materials (subatomic particles, atoms, molecules, gunk) that are contributing biofunctionally to my activities now are not identical to the materials that contributed biofunctionally to my activities last year. Hylomorphists offer a similar analysis of statement (2). It says that the materials that contributed biofunctionally to my activities on July 3, 1997 are the very same materials that are now contributing biofunctionally to the activities of the tree. Finally, hylomorphists can clarify the meaning of statement (3) by describing how the materials that contributed biofunctionally to my activities on July 3, 1997 are now contributing biofunctionally to the activities of many different living things. This is the sense in which those materials are dispersed throughout the biosphere: they are now parts of other living things.

Does it make a difference to the semantics of these expressions if the matter in question is atomic or gunky? It is hard to see how it could. The reason is that in this context, the term 'matter,' like the term 'materials,' is being used in a way that is neutral about the nature of that matter. If the matter is gunky, then statement (1) says that the gunky matter that is contributing biofunctionally to my activities now is not identical to the gunky matter that contributed biofunctionally to my activities last year; statement (2) says that the gunky matter that contributed biofunctionally to my activities on July 3, 1997 is the very same gunky matter that is now contributing biofunctionally to the activities of the tree; and statement (3) says that the gunky matter that contributed biofunctionally to my activities on July 3, 1997 is now

contributing biofunctionally to the activities of many different living things. This seems to be a plausible semantics for expressions like 'the matter composing me at *t*.' Consequently, it is not implausible to conclude that hylomorphists do not face a serious worry about the possibility of gunk.

7.4 The Denial

A second objection to van Inwagen's view of composition concerns the Denial. Recall that according to van Inwagen, the only material beings are living things and mereological simples (Section 6.2). Since many objects in a commonsense ontology, such as tables, mountains, and planets, are neither living things nor mereological simples, van Inwagen's view implies that these things do not exist. But, say critics, this implication is absurd.

We've seen that hylomorphism implies a version of the Denial. Its ontology includes fundamental physical materials, structured individuals, and their biofunctional parts. It denies the existence of artifacts, and is likely to deny the existence of a wide variety of natural bodies, for reasons we'll consider momentarily. Aristotle's hylomorphism also implies a version of the Denial, although this point is obscured by his terminology. Unlike most contemporary metaphysicians, Aristotle is willing to use the term 'being' equivocally. He is thus willing to say on occasion that there are (in some sense) artifacts:

> Some things are said to be by bringing together their matter; for example, some things are said to be by mixing, such as honeywater; others by tying, such as a bundle; others by gluing, such as a book; others by nailing, such as a box; others by more than one of these; others by position, such as a threshold or a lintel (for these exist by being positioned in some way); others by time, such as dinner and breakfast; others by location, such as the winds; others by their perceptible features, such as hardness and softness, thickness and thinness, dryness and wetness—either some of these or all of them, and in general by their excess or deficiency. Clearly, then, 'is' is said in just as many ways. Something is a threshold, for instance, because it lies in this position, and being a threshold means having this position; similarly, being ice means having solidified in this way. The being of some things will be defined by all of these things; some by being mixed, some blended, some tied together, some solidified, and some, such as a hand or foot, by having the other differentiae. (*Metaphysics* 1042b15–31)

But Aristotle's equivocation on 'being' is not unsystematic; it is what he calls an equivocation *pros hen*—literally 'toward one.' The various equivocal senses of 'being,' he tells us, are united by reference to a univocal core (*Metaphysics* 1003a33–b10). Substances are beings in the strictest sense; to say that there are beings in any other sense entails in one way or other that there are substances. If anything in Aristotle's metaphysics maps onto our univocal notion of being, it is the sense of 'being' that applies to substances. According to Aristotle, however, substances exist by nature, and the things that exist by nature include the elements (earth, air, fire, and water),

plants and animals, and their parts, but they do not include artifacts (*Physics* 192b8–12; *Metaphysics* 1042a8–10).[6]

If terms like 'being' and 'exists' are equivocal, then hylomorphists can say that there are artifacts, albeit in a sense of 'are' different from the sense in which they say that there are fundamental physical materials or living things. In what follows, however, I will assume that 'being,' 'exists,' and related terms are used univocally. That means hylomorphists must look to defend the Denial.

One argument in favor of the Denial appeals to redundant causal powers (van Inwagen 1990: 122; Merricks 2001a: chapter 3). Suppose that I want to explain a particular state of affairs: I want to know how a window on my house got shattered. The explanation is one that in our pedestrian dealings we would express by saying that Eleanor threw a baseball that shattered the window. This pedestrian explanation posits a baseball and a window. Because defenders of the Denial reject the existence of artifacts, they will look to paraphrase this explanation in a way that makes no reference to windows or baseballs. They will say, for instance, that there are particles $a_1, a_2, \ldots, a_n$ and particles $b_1, b_2, \ldots, b_n$ (I am assuming here for simplicity that some type of atomism is true), and that Eleanor accelerated the *b*s toward the *a*s and displaced some of the latter. According to defenders of the Denial, this paraphrase does all the explanatory work that needs to be done in this context; it provides an exhaustive explanation of why some of the *a*s ended up on the ground, a state of affairs described in our pedestrian dealings as a window on my house being shattered. If defenders of the Denial are right, then explaining this state of affairs does not require us to posit a window and a baseball in addition to the *a*s and the *b*s, and if that is the case, there is good reason to think that baseballs and windows do not exist. To understand why, let us suppose for the sake of argument that the *a*s compose a window, and the *b*s compose a baseball. The Eleatic Principle endorsed in Section 2.2 implies that the only things that exist are ones that play some sort of causal role. This implies that the only individuals that exist are ones that can enter into causal relations. Consequently, if the baseball and the window exist, they must be capable of entering into causal relations, they must have causal powers.

Suppose, then, that the baseball and the window have causal powers. In that case, either (a) the powers they have must be powers to produce exactly the same effects as their proper parts, or else (b) the powers they have must be powers to produce effects different from the effects produced by their proper parts. If (a) is true, then the effects the window and the baseball produce must be overdetermined, for if (a) is true, then the window and the baseball will have only powers to produce exactly the same effects as those produced by the *a*s and the *b*s: the effect of the baseball striking the window will be exactly the same as the effect of the *b*s displacing some of the *a*s. That

[6] He thinks parts are of two sorts: nonuniform parts (*ta anomoiomerē*), which are organs like the heart, and uniform parts (*ta homoiomerē*), the kinds of things we would likely call 'tissues': heart tissue, brain tissue, and so on (*Parts of the Animals* 647a4ff).

effect will be produced by two causal pairs: the window and baseball exercising their powers, on the one hand, and the *a*s and *b*s exercising their powers, on the other. The exercise of either pair of powers by itself will be sufficient to produce the effect, for by hypothesis each pair has the power to produce exactly the same effect. Presumably, moreover, at least one of the pairs is sufficient to produce the effect. In that case, however, both must be.

The upshot: if (a) is true, then artifacts and the materials that compose them will overdetermine their effects. But systematic overdetermination of this sort is problematic. It is a position that fails to match many of our epistemic intuitions, and that is, in general, a mark of uneconomical theorizing (Merricks 2001a: 66–7). In addition, overdetermination has absurd implications in many cases. Consider an example from the philosophy of mind. Suppose your actions are overdetermined by your mental states together with events involving the proper parts composing your nervous system; events of each sort are sufficient causes of your actions. If each is sufficient, however, then neither by itself is necessary. Your actions could have been caused by the events in your nervous system alone without your mental states, and they could have been caused by your mental states alone without the events in your nervous system. But both of these results seem absurd. On the one hand, it seems absurd to suppose that your actions could have occurred without your beliefs or desires. Simply to qualify as actions events need mental causes. On the other hand, it seems absurd to suppose that your actions could have occurred without the events in your nervous system—the events that were responsible for triggering the muscular contractions in your limbs. Because overdetermination has absurd results like these in many different cases, defenders of option (a) would have to construct an account that distinguished in a principled way the problematic cases from the rest. But constructing any such account is bound to be difficult.

Suppose, then, that (b) is true instead, that the window and the baseball have powers to produce effects different from the effects produced by the *a*s and the *b*s. In that case, those causal powers must be emergent: they must be powers to produce effects that cannot be exhaustively explained by appeal to the *a*s and the *b*s. Hylomorphists do not reject the existence of emergent powers; on the contrary, they take such powers to be a hallmark of structured individuals (Section 6.3). Here, however, we are assuming that the window and the baseball do not have any emergent powers. Our goal, recall, is to defend the premise that if an explanation that appeals to the *a*s and *b*s alone is exhaustive, then there is good reason to think that neither windows nor baseballs exist. If an explanation that appeals to the *a*s and the *b*s is exhaustive as we've assumed, then there are no emergent causal powers: there are no effects other than those that are explained by the *b*s displacing some of the *a*s.

There is thus good reason to deny both (a) and (b), but if we deny both, then we must deny that the window and the baseball have causal powers, and if we deny that, then by the Eleatic principle we must deny that the window and the baseball exist.

In response, opponents of the Denial can target one of the argument's premises. They can embrace the systematic overdetermination of effects by artifacts and their proper parts, or they can accept that artifacts have emergent causal powers, or they can claim that some individuals entirely lack causal powers. But if these claims are as unpalatable as they seem, the upshot is that artifacts do not exist. There is good reason to accept the Denial in their case at least. Something analogous seems true of mountains, planets, and many other natural bodies. If they existed, they would do nothing other than what their proper parts do collectively. If we accept the Eleatic Principle and reject systematic overdetermination, then we have good reason to think that they do not exist.

Hylomorphists do not deny a priori that anything other than fundamental physical materials and living things exist. We've seen that unlike van Inwagen's view, theirs is compatible with the discovery that some structured individuals are nonliving. Those individuals, if they existed, would have emergent causal powers, powers that could not be exhaustively described and explained by appeal to the materials composing them. Whether there are nonliving things like this is something that hylomorphists take to be an open empirical possibility. This follows from their commitment to ontological naturalism (Section 1.2). If our best empirical descriptions, explanations, and methods postulate entities of kind *K*, then we have good prima facie reason to think that *K*s exist. If, for instance, our best empirical descriptions, explanations, and methods indicate that strands of DNA have causal powers that cannot be exhaustively described and explained by appeal to the fundamental physical materials that compose them, if they instead have emergent powers, then there is good reason to think that those strands exist just as organisms do.

7.5 The Vagueness Argument

Another objection to van Inwagen purports to show not just that his view of composition is false, but that any restrictivist view of composition must be false. This is the vagueness argument (Lewis 1986a: 211–13; Sider 2001: 120–34). Here is one version of it:

(1) If composition is restricted, then it is sometimes indeterminate whether a composite object exists or not.
(2) If it is sometimes indeterminate whether a composite object exists or not, then some logical expressions are vague.
(3) No logical expressions are vague.

Therefore, it is never indeterminate whether a composite object exists or not.
Therefore, composition is not restricted.

There are at least two arguments supporting Premise (1). The first is advanced by David Lewis. He argues that any principle of restricted composition that is based on

our ordinary intuitions about what composite objects exist must be vague. But if such a principle is vague, then composition must be vague as well:

> To restrict composition in accordance with our intuitions would require a vague restriction ... We are happy enough with mereological sums of things that contrast with their surroundings more than they do with one another; and that are adjacent, stick together, and act jointly. We are more reluctant to affirm the existence of mereological sums of things that are disparate and scattered and go their separate ways ... It's not on to say that somewhere we get just enough contrast with the surroundings, just enough cohesion ... to cross the threshold and permit composition to take place, though if the candidate class had been just a little worse it would have remained sumless. But if composition obeys a vague restriction, then it must sometimes be a vague matter whether composition takes place or not. (Lewis 1986a: 211–12)

The second argument for (1) is advanced by Ted Sider (2001: 122). If composition is restricted, he says, then it is possible to construct a sorites series involving possible cases of composition. Imagine a continuum of possible cases involving the objects a_1, $a_2, \ldots, a_n$. At one end of the continuum are cases in which the *a*s compose nothing. At the other end are cases in which the *a*s compose something. According to Sider, one of two things must be the case: either there must a sharp cutoff on the continuum or not. If there is a sharp cutoff, then there will be an *i*th case on the continuum at which the *a*s compose something even though they compose nothing at the *i*th-minus-1 case. If, on the other hand, there is not a sharp cutoff, then there will be no such case; there will instead be a number of cases in which it is indeterminate whether or not the *a*s compose something; that is, composition will be vague. It is implausible to suppose that there could be a sharp cutoff on a continuum like this, says Sider. The reason is that admitting a sharp cutoff at any point on the continuum seems arbitrary. If the cases on the continuum are very similar to each other, and if each case is as similar to adjacent cases as every other case, there seems no principled reason to admit a sharp cutoff at one point on the continuum rather than another. But if there is no nonarbitrary cutoff, then it seems that composition must be vague. Sider concludes that if composition is restricted, it must sometimes be vague whether a composite object exists or not.

Consider now Premise (2). If objects *a* and *b* compose something, then there are not two objects, but three: *a*, *b*, and the object they compose. Suppose, however, that it is indeterminate whether *a* and *b* compose something. In that case, it is indeterminate whether there are two objects or three; it is indeterminate, in other words, how many objects exist. Sider assumes that if we are to locate the source of this indeterminacy anywhere, it must be in language, since he follows David Lewis in endorsing a linguistic theory of vagueness:

> The only intelligible account of vagueness locates it in our thought and language. The reason it's vague where the outback begins is not that there's this thing, the outback, with imprecise borders; rather there are many things, with different borders, and nobody has been fool enough

to try to enforce a choice of one of them as the official referent of the word 'outback'. Vagueness is semantic indecision. (Lewis 1986a: 212)

Given this account of vagueness, it follows that indeterminacy about composition will be expressible using what Sider calls *numerical sentences*, ones which assert that there are exactly *n* concrete objects. A numerical sentence asserting that there are exactly two concrete objects would be '$\exists x \exists y\ [Cx\ \&\ Cy\ \&\ x \neq y\ \&\ \forall z(Cz \rightarrow [x=z \vee y=z])]$.' Since the indeterminacy can be expressed using sentences like these, if we are to locate the source of the indeterminacy in language, we must locate it in some vague predicate or term—either in the predicate '*C*,' which expresses concreteness, or else in one of the logical expressions: $\exists$, $\forall$, $=$, &, $\vee$, $\rightarrow$. But we cannot locate it in '*C*,' says Sider. The reason is that we can stipulate that concrete objects are ones that do not belong to any of the following categories: sets and classes, numbers, properties and relations, universals and tropes, and possible worlds and situations. Since it is not indeterminate whether something belongs to one of these categories, the predicate '*C*' cannot be vague. We must therefore locate the indeterminacy in one of the logical expressions.

According to Premise (3), however, it is implausible to suppose that any of the logical expressions could be vague. The reason is that according to the linguistic theory of vagueness, expressions are vague on account of semantic indecision: every vague expression admits of multiple precisifications, multiple ways of giving the expression a precise meaning, yet without a way of determining which precise meaning to endorse in fact. The operators &, $\vee$, and $\rightarrow$ for their part do not have multiple precisifications, and hence cannot be vague. Likewise, there are good reasons to think that the identity predicate, $=$, cannot be vague (Evans 1978; Salmon 1981: 243–5). That leaves the quantifiers. In numerical sentences, these are unrestricted, but it is implausible, Sider argues, to suppose that an unrestricted quantifier could admit of multiple precisifications. Suppose that $\forall_1$ and $\forall_2$ correspond to two different precisifications of $\forall$. In that case, there would have to be an *x* which belonged to the extension of one, but not to the extension of the other. Suppose, then, that *a* belongs to the extension of $\forall_1$ but not the extension of $\forall_2$. In that case, $\forall_2$ cannot be a precisification of the unrestricted quantifier, for its extension is clearly restricted: it excludes *a*. It is not possible, therefore, to locate the indeterminacy in any of the logical expressions, and since it is not possible to locate it in the predicate '*C*' either, it follows that it can never be indeterminate whether or not some composite object exists. But if that can never be indeterminate, then composition cannot be restricted.

There are several ways defenders of restricted composition can respond to the vagueness argument. I will focus on a line of response due to Trenton Merricks (2005). It targets Premise (1). First, Merricks (2005: 619–23) argues that Lewis' argument for (1) fails because it attacks a straw man. Lewis tacitly assumes that a principle of restricted composition must be based on our ordinary intuitions about what composite objects exist. But this assumption is unfounded. A commitment to

accommodating ordinary intuitions about composition is neither motivated by nor implied by restrictivism. This point is illustrated by van Inwagen's view, which is a well-motivated restrictivist view that defies some ordinary intuitions about composition.

Second, Merricks argues that Sider's argument for (1) fails if there are emergent powers. Consider again the composition continuum discussed earlier. It is a series of possible cases involving the objects $a_1, a_2, \ldots, a_n$. At one end of the continuum are cases in which the *a*s compose nothing, and at the other end are cases in which they compose something. One of the premises in Sider's argument for (1) is that it is implausible to suppose there could be a sharp cutoff on the continuum since admitting a sharp cutoff at any point would be arbitrary. Merricks (2005, 2007) argues, however, that some continua admit nonarbitrary cutoffs. There is, for instance, a sharp cutoff between the *i*th straw that finally breaks the camel's back and the *i*th-minus-1 straw that doesn't. The reason is that the *i*th straw results in something completely different from the mere addition of another straw. In this sense, a straw-adding continuum is unlike a baldness continuum, for on a baldness continuum no new effect results from the addition or subtraction of a single hair. Suppose, then, that composition results in something completely new, an object with emergent properties not had by the materials out of which it is composed. If it is never indeterminate whether an object has those properties, then it is not implausible that there should be a sharp cutoff on the composition continuum. The *i*th case in the continuum at which a composite object first appears will be the case at which the emergent properties first appear. Merricks illustrates the idea by imagining that composite objects essentially and continuously emit a whistling sound which is not emitted by the materials composing them. If it is never indeterminate whether the whistling occurs, then the *i*th case in the continuum at which a composite object first appears will be the case at which the whistling is first heard.

If composite objects have emergent properties, therefore, then Sider's premise is false, and his argument for Premise (1) of the vagueness argument fails. Moreover, if composite objects have emergent properties, then Premise (1) of the vagueness argument is itself false, for if composite objects have emergent properties, then it is possible for composition to be restricted without it ever being indeterminate whether a composite object exists. To see this, simply imagine that composition occurs only when some things are configured in a way that enables whistling, and that not everything is configured this way. If we continue to assume that composites essentially and continuously whistle, and that it is never indeterminate whether whistling occurs, then we have the model of a case in which composition is restricted, but in which it is never indeterminate whether a composite object exists.

Are there such cases in fact? If hylomorphists are right, then there are, for according to hylomorphists, composite individuals have emergent properties. In particular, they have individual-making structures, powers to configure the materials that compose them: to grow, to develop, to maintain themselves counter-entropically,

and to do many other things that it is the business of biology and other empirical disciplines to describe. If hylomorphists are right, then Premise (1) of the vagueness argument is false. Composition on the hylomorphic view is restricted to structured individuals, yet it need never be indeterminate whether a structured individual exists since structured individuals have properties that are not had by the materials that compose them. Merricks' objection to the vagueness argument thus dovetails with the hylomorphic view.

7.6 Conclusion

The goal of this chapter was to defend the hylomorphic view of composition against some objections. First, van Inwagen argues that there are no eyes, hands, hearts, or other biofunctional parts on the grounds that such parts are arbitrary. Hylomorphists argue that such parts are not arbitrary. Ontological naturalism, which was adopted as a working hypothesis in Section 1.2, implies that empirical adequacy is an important criterion for choosing a principle of part identity and individuation: if principle *P1* does a better job enabling us to describe and explain the behavior of *K*s than principle *P2*, then we should accept *P1* over *P2*. Actual work in biology and biological subdisciplines such as neuroscience, moreover, suggests that a principle of part identity and individuation that postulates biofunctional parts such as eyes, hands, and hearts is more adequate for describing and explaining real living behavior than other principles. This provides grounds for choosing it in favor of other principles.

Second, Trenton Merricks and Eric Olson argue that if biofunctional parts such as brains exist, then we confront the problem of too many thinkers. If my brain and I both exist, then we must both be rational consciousness beings if either of us is. Yet, it is absurd to suppose that there could be two rational, conscious beings located exactly where I am; hence, my brain must not exist, and the same is true by extension of any biofunctional part. Hylomorphists respond that the argument's premise is false: it does not follow that my brain must be a rational, conscious being if I am.

Third, several objections have been advanced against van Inwagen's account of composition based on the possibility of atomless gunk and the Denial, the claim that artifacts and natural bodies do not exist. Since the hylomorphic account is similar to van Inwagen's, there is a worry that the same objections can be advanced against it. In fact, however, this is not the case. Hylomorphists have ways of handling these objections.

Finally, hylomorphists have a response to the vagueness argument. The latter purports to show that composition cannot be restricted. The argument fails if there are emergent properties. Since hylomorphism implies that there are emergent properties, the vagueness argument fails to show that hylomorphic composition cannot be restricted.

8

Structured Activities and Embodiment

8.1 Activity-Making Structures

The discussion of structure has so far focused on the kinds of structures that make individuals what they are: we called them *individual-making structures*. But individual-making structures are not the only structures that exist according to hylomorphists. The activities in which structured individuals engage have structures as well. Since activities are not individuals, the structures they have are not individual-making structures; they are instead *activity-making structures*.

The idea that there are activity-making structures is based on the observation that the activities of structured individuals involve coordinated manifestations of the powers of their parts. Consider the activities we humans engage in: walking, crawling, talking, singing, dancing, reaching, grasping, running, jumping, throwing, riding, driving, typing, swimming, reading, eating, breathing, and so on. Each of these activities involves some of our parts manifesting their powers in an organized or coordinated way. When we perform these activities, we impose an organization or order on the ways those parts manifest their powers. In doing so, we do not bring into existence new individuals, but rather activities, which are events: my walking, his crawling, her talking, and so on. What is true of these activities, moreover, is also true of activities that have to be discovered through scientific investigation. At one level, for instance, we engage in metabolic activities—processes whereby we tap the energy stored in the chemical bonds of materials we've acquired from the environment. The Krebs cycle is an example (Figure 8.1). It is a metabolic process that begins with the entry of pyruvate molecules into the mitochondria of cells. Since pyruvate molecules are charged, they are shuttled through the mitochondrial membrane with the help of transport proteins. Once the molecules enter the mitochondrial matrix, enzymes within the matrix catalyze reactions which yield carbon dioxide, and the molecules NADH and acetyl coenzyme A. The last of these is composed of an acetyl group attached to coenzyme A by an unstable bond. Because the bond is unstable, it is easily broken and the acetyl group binds itself to an oxaloacetate molecule to form citrate. The citrate is oxidized yet further in a series of reactions that yield ATP, a molecule with a weakly bonded phosphate group. When that bond is broken, the reaction yields energy for the cell to do work.

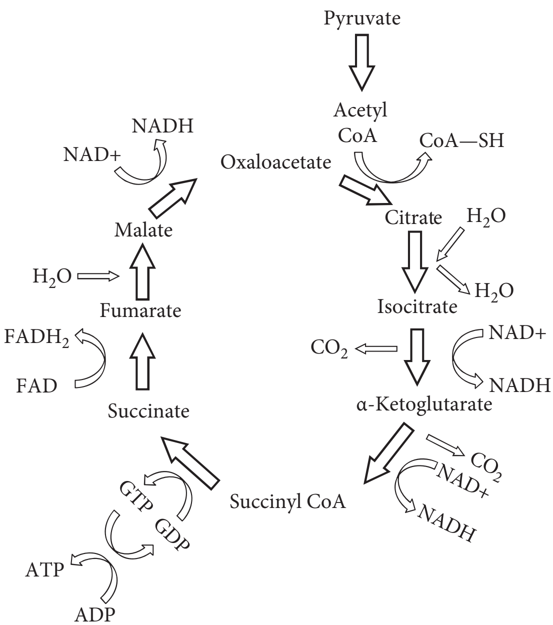


Figure 8.1 The Krebs cycle.

The Krebs cycle illustrates an important point first introduced in Chapter 1: some structures are dynamic. They are not relatively stable spatial relations among an individual's parts. They instead have a temporal dimension: they comprise sequences of changes over time. They often have a conditional dimension as well: they involve different kinds of changes under different kinds of conditions. For that reason, we often represent them using flowcharts like Figure 8.1.

The Krebs cycle, as well as walking, crawling, talking, and the other activities mentioned above, illustrate two important points about the activities of structured individuals. First, those activities involve coordinated manifestations of an individual's parts. The Krebs cycle is not a random assortment of energy transformations, it is an ordered sequence of occurrences involving the molecular parts of a cell: transport proteins, enzymes, pyruvate, citrate, and other molecules. Likewise, swimming, throwing, and walking are not randomly assorted limb movements; they are instead coordinated sequences of occurrences involving our limbs.

Second, the coordinated manifestation of the powers of something's parts is the very thing that qualifies those parts as parts. Recall that individuals qualify as parts of more inclusive wholes on account of contributing biofunctionally to the activities of those wholes (Section 6.4). In the case of the Krebs cycle, various molecules qualify as parts of a cell on account of playing roles in a coordinated sequence of events that supply the cell with energy in the form of ATP. It is because the molecules involved in the Krebs cycle manifest their powers in this coordinated manner that they contribute biofunctionally to the cell's activities, and it is because they contribute

biofunctionally to its activities that they qualify as parts of the cell. The same is true mutatis mutandis of the cell itself. Its activities and those of other cells are coordinated in ways that contribute biofunctionally to more inclusive tissue or organ systems. The activities of the cells thereby qualify them as parts of those more inclusive systems, and the activities of the systems themselves are coordinated in ways that contribute biofunctionally to the activities of the organism as a whole—contributions that qualify them as parts of the whole.

According to hylomorphists, the coordinated manifestation of the powers we find in the Krebs cycle and other activities is a species of structuring, another manifestation of the power organisms have to impose order on materials. When I throw a baseball, swim, or play a musical instrument, I manifest my power to structure the way my biofunctional parts manifest their powers. My parts needn't manifest their powers in this way. It is possible for my neurons to fire or my muscles to contract in ways that do not compose an activity of throwing a baseball, swimming, or playing an instrument. Fatigue, injury, insufficient training, and many other factors can result in uncoordinated manifestations of the powers of my biofunctional parts. But when I succeed in throwing, swimming, or playing, I succeed in imposing a structure on the way my parts manifest their powers: I structure their manifestations throwing-, swimming-, and playing-wise, respectively.

These examples illustrate a further point: a large number of activities in which things like us engage involve the coordinated manifestation of the powers not just of our parts, but of surrounding materials as well. Throwing a baseball, swimming, playing an instrument, and many other activities have environmental requirements. Throwing a baseball requires a baseball, physical materials spatially arranged baseball-wise. Those materials, moreover, must be properly disposed; they must be free to travel when I release them, for instance: they cannot be taped or glued to my hand. Likewise, swimming requires being in water, and playing an instrument requires an instrument (properly disposed). Many activities have social requirements as well. Playing chess requires not just the coordinated movements of one's limbs and the availability of moveable chess pieces on a board, it also requires that my opponent and I know and follow the rules of the game. Something analogous is true of voting, bidding, buying, promising, and marrying. All of them have social requirements in addition to environmental ones.

Suppose, then, that *a* is a structured individual that has the power to engage in activity *A*. We can say in general that *a* manifests its power to *A* exactly if *a*'s parts and the surrounding materials manifest their powers in the right kind of way. More precisely, *a* engages in *A*-ing exactly if the manifestations of the powers of its parts and the surrounding materials are structured *A*-wise, where it is an empirical undertaking to determine what it takes for something to be structured *A*-wise.

Activity-making structures have characteristics similar to those of the individual-making structures discussed in Section 6.1. First, just as structured individuals configure the materials that compose them, they also configure the activities of

their parts, and in many cases the activities of surrounding materials as well. When I throw a baseball, I coordinate the activities of my muscles, nerves, and things in the environment, such as the positions and properties of physical materials that are arranged baseball-wise. I cause all of these things to manifest their powers in the way it takes for me to throw a baseball.

Second, activity-making structures are particulars. My throwing a baseball at t_1 is an event in which I impose order on the way my parts and various surrounding materials manifest their powers. That imposing of order is a nonrepeatable particular which is nontransferable in the sense described in Section 3.2: it belongs essentially to my throwing a baseball at t_1. My throwing a baseball at t_2 might exactly resemble my throwing a baseball at t_1, but it is a different event. The two events are different (if exactly similar) impositions of order.

Third, like individual-making structures, activity-making structures are manifestations of powers and also powers for further manifestations. Throwing a baseball is the manifestation of a power that I have to coordinate the activities of my parts and the surrounding physical materials, but it is also a power to make an out in a baseball game or to knock down a stack of cans at the county fair. The same is true mutatis mutandis of the latter activities: each is a manifestation of my power to throw a baseball, but each is also a power for further manifestations. The first is a power to contribute to my team's victory and please my teammates; the second is a power to win a teddy bear and please my daughter.

Fourth, activity-making structures have the same directedness that all powers do. Just as my power to throw a baseball is directed toward its manifestation, likewise that manifestation, the activity of throwing a baseball, is a power that is directed toward manifestations of its own, such as the activity of making an out or of knocking down cans.

It's important to note that in the case of animals like us, the ends toward which our powers are directed are sometimes ends that we choose.[1] We manifest our power to make choices in the ends toward which some of our activities are directed. Intentional action is just one species of the end-directedness that characterizes activities in general on the hylomorphic view. That end-directedness, moreover, is just a feature of activities being powers, a feature implied by the general theory of properties and powers defended in Chapters 2–5.

Fifth, activity-making structures confer whatever powers they do necessarily. It is metaphysically impossible for my throwing not to confer on me the power to make an out in a baseball game or to knock down cans at the county fair. Someone might object that we can easily conceive of situations in which the activity of throwing does not confer these powers on me—situations, for instance, in which I inhabit a culture that doesn't have baseball games or county fairs. In response, hylomorphists can

[1] Hylomorphists are not committed to any particular analysis of choice—whether, say, it is to be understood in a libertarian or compatibilist sense.

argue that in such situations I would still have the powers to make an out and to knock down cans, I would just never be able to manifest them. By analogy, salt crystals on a waterless planet have the power to dissolve in water; they simply cannot manifest it. Critics might persist with yet other examples that supposedly illustrate the conceivability of situations in which an activity does not confer the powers it actually does. We have seen already how hylomorphists look to deal with these cases (Sections 4.3 and 6.1). Either they can deny the conceivability of worlds in which my throwing does not confer on me the power to make an out or to knock down cans, or else they can deny that the kind of conceivability we achieve in these cases is a reliable guide to possibility.

With the foregoing points in mind, consider again the theoretical roles we expect structure to play (Sections 1.1 and 6.1):

Structure matters: it operates as an irreducible ontological principle, one that accounts at least in part for what things essentially are.

Structure makes a difference: it operates as an irreducible explanatory principle, one that accounts at least in part for what things can do, the powers they have.

Structure counts: it explains the unity of composite things, including the persistence of one and the same living individual through the dynamic influx and efflux of matter and energy that characterize many of its interactions with the wider world.

Structure minds: it provides us with resources for understanding the place of mental phenomena within the natural world.

Based on what's been said, it should be evident how activity-making structures matter. Structured activities have their activity-making structures essentially. When I throw a baseball (physical materials that are spatially arranged baseball-wise), my movements are essentially coordinated in the way it takes for me to engage in that activity. The event, my throwing a baseball at time *t*, exists exactly if my parts and the surrounding materials are coordinated throwing-a-baseball-wise—exactly if they have that activity-making structure. If my parts and the surrounding materials had some other activity-making structure, if they were configured playing-an-instrument-wise, they would compose an entirely different activity. Activity-making structures are ontological principles, therefore: they matter; they account in part for what the activities they configure essentially are.

Activity-making structures also make a difference. The reason is that they are powers. If an individual engages in a structured activity, that activity confers on the individual powers that it would not otherwise have. If I am engaged in the activity of throwing a baseball—if my parts and the surrounding materials are structured throwing-a-baseball-wise—then I have the power to make an out or to knock down cans, and these are powers I would not possess if my behavior were not structured this way. Activity-making structures thus make a difference.

Activity-making structures also count: they confer unity on diverse events in something analogous to the way individual-making structures confer unity on the

physical materials that compose structured individuals. My parts and the surrounding materials needn't manifest their powers throwing-a-baseball-wise. The very same muscle fibers that contract in my shoulder when I throw a baseball might also contract when I experience an uncontrolled muscle spasm, or when a physical therapist stimulates them electrically. What unifies the contractions of the muscle fibers, what coordinates the manifestations of their powers, is what I do when I try to make an out, or try to knock down cans, or try to accomplish whatever I try to do when I throw a baseball. In undertaking these activities, I impose a unified order on the way my parts and the surrounding materials manifest their powers.

8.2 Activity Composition

Recall that when it came to individual-making structures, we described the unifying role of structure in terms of composition (Section 6.3–6.4). On the hylomorphic view, composition occurs when and only when an individual configures or structures materials: there is a y such that the xs compose y if and only if y is an individual that structures the xs. This claim was intended to apply only to individuals, but it is possible to articulate an analogous notion of composition for activities. Just as physical materials compose an individual exactly if the right kind of individual-making structure is imposed on them, likewise various events compose an activity exactly if the right kind of activity-making structure is imposed on them.

Here is one attempt to define a notion of activity composition: Suppose that a is a structured individual with the power to engage in activity A. Suppose, moreover, that b_1, $b_2, \ldots,$ b_n are individuals, and that a subset of the bs are proper parts of a. Suppose, finally, that each b_i has the power to engage in an activity, A_i: b_1 has the power to A_1, b_2 has the power to A_2, and so on. We can then define a notion of activity composition as follows:

Activity composition: Necessarily, a's A-ing at t is composed of b_1's A_1-ing at t_1, b_2's A_2-ing at $t_2, \ldots$ and b_n's A_n-ing at t_n exactly if b_1's A_1-ing at t_1, b_2's A_2-ing at $t_2, \ldots$ and b_n's A_n-ing at t_n are structured A-wise, where time t includes the times t_1, $t_2, \ldots, t_n$.

Five notes are in order about this definition of activity composition. First, the kinds of events that count as composite activities are fairly limited. The definition requires that the composing events include manifestations of the powers of something's proper parts. According to hylomorphists, however, the only individuals that have proper parts are structured individuals. Consequently, the only composite activities will be the activities of structured individuals, paradigmatically the activities of living things.

Second, the bs whose activities compose a's A-ing must include some of a's proper parts, but the bs can include—and in some cases must include—individuals that are not among a's proper parts. My throwing a baseball requires that there be individuals

in the surrounding environment that are spatially arranged baseball-wise, individuals which are not parts of me, but which belong to the surrounding environment.

Third, not all of an individual's powers are relevant to its role in composing an activity. A proper part of *a* might have powers that contribute in no way to *a*'s *A*-ing. Just as the color of a gear is not what enables it to mesh with the rest of a mechanism in which it is embedded, but only its size and shape, likewise only some of the powers of an individual b_i might be the powers that enable it to contribute to *a*'s *A*-ing.

Fourth, given reasonable assumptions, activity composition implies that the behavior of structured individuals never violates the laws governing their fundamental physical constituents. According to hylomorphism, the activities of structured wholes are composed of the structured manifestations of the powers of their lower-level constituents and surrounding materials. If those constituents or materials were to lose their powers, or were to become incapable of manifesting them, they would become incapable of composing the activities of structured wholes. Those activities depend on lower-level items retaining and manifesting the powers they have. By analogy, it is only because bricks and timbers retain their shapes under compression that they can be recruited as components of buildings. Likewise, it is only because lower-level materials retain their distinctive powers that structured individuals can recruit them as components for their own activities.

Finally, this definition of activity composition implies a way of understanding the idea that a structured individual has its powers *in virtue of* having its parts; in particular, it implies that a structured individual has the power to engage in various activities because it has parts, and those parts form a subset of the individuals with powers whose coordinated manifestations compose its activities.[2] In this sense, we can say that an individual is empowered in virtue of having the parts it has. It will be helpful to introduce some more terminology to express this idea.

Let us say that the parts of a structured individual *embody* its powers, or that its powers are *embodied in* its parts. My visual system embodies my power to see; your circulatory system embodies your power to bring oxygenated blood to various parts of yourself; Gabriel's nervous system embodies his power to coordinate the movements of his limbs, and so on. Let *a* be a structured individual with the power to engage in activity *A*, and let $b_1, b_2, \ldots, b_n$ be individuals, a subset of which are proper parts of *a*. Suppose now that *a*'s *A*-ing at a time would be composed of the *b*s manifesting their powers *A*-wise at that time. In that case, let us say that *a*'s power to *A* is embodied in those *b*s which are proper parts of *a*: those *b*s embody *a*'s power to *A*.

[2] Some authors use the expression 'in virtue of' to designate the grounding relation. It should be evident that I am using the expression in a different sense. My sense of 'in virtue of' concerns the essential embodiment of something's powers, which is a notion different from that of grounding. Grounding plays a significant role in the hylomorphic theories defended by Rea (2011) and Koons (2014).

Note that this definition of embodiment implies that an individual's powers are embodied in its parts alone, not in its parts plus surrounding materials. Even though the latter are in some cases necessary for an individual to manifest its powers, the definition limits embodiment to the parts of individuals. The reason for this limitation is that in a wide variety of cases individuals remain empowered even in the absence of surrounding materials in the right condition. I retain the power to throw a baseball even if no baseballs are available. Consequently, if a structured individual's powers are essentially embodied, then its powers cannot be embodied in surrounding materials.

8.3 The Embodiment Thesis

Are all of the powers of structured individuals essentially embodied? Let us call the claim that they are the *embodiment thesis*. Many hylomorphists of the past have denied the embodiment thesis. Aristotle himself appears to deny it in *De Anima* III.4. There he appears to argue that understanding or *nous*, the power to grasp the essences of things, has no organ and is in general unmixed (*amigēs*) with a body:

> We must now inquire about the part of soul by which it knows and thinks—whether it is separable [from bodily things], or not separable in magnitude but only in definition... If understanding (*nous*) is like perceiving... it must be capable of receiving the form [of what is understood] without being it, and it must be related to what is understood as the perceptive part [of soul] is related to what is perceived... Therefore, since everything is understood, it must be unmixed [with anything bodily]... for the intrusion of anything foreign would hinder and obstruct it... It thus has no nature except this: that it is potential. Hence, the part of the soul called 'understanding'... is actually none of the things which exist before it understands. It is thus reasonable that it should not be mixed with a body, for then it would acquire a quality of the body, either cold or hot, or even have an organ like the perceptive part does, but in fact it has none. (429a10–27)

Arguments like Aristotle's pose a prima facie challenge to the embodiment thesis. What can defenders of the thesis say in response?

First, it's important to recognize that a commitment to the essential embodiment of our capacities is the default position for the hylomorphic theory I've described. It should be relatively uncontroversial that we cannot do most of the things we do without some of our parts manifesting their powers in coordinated ways. Examples include walking, crawling, talking, singing, dancing, reaching, grasping, running, jumping, throwing, riding, driving, typing, swimming, reading, eating, breathing, seeing, hearing, smelling, tasting, and touching. According to the hylomorphic theory I've described, these activities are essentially composite: they are essentially composed of the coordinated manifestations of the powers of our parts and surrounding materials. It is impossible, not just nomologically, but metaphysically, for me to engage in the activity of throwing a baseball unless my parts and the

surrounding materials manifest their powers in a coordinated way. Given the range of examples that conform to this model, it seems reasonable to conclude that the same is true of all our activities, including thinking or understanding.

A commitment to essential embodiment is the default position not just for the kind of hylomorphic theory I've described—one committed to an account of activity composition. Rather, essential embodiment would appear to be the default position for any hylomorphic theory, including Aristotle's. Aristotle originally articulates his hylomorphic framework in order to defend the existence of change against the arguments of Eleatic philosophers such as Parmenides and Melissus. In that framework, every change involves some underlying thing (the matter) taking on some characteristic (a form). A statue, for instance, comes to be by some quantity of bronze becoming statue-shaped. When Aristotle looks to apply this framework to psychological capacities, it is not surprising that he tells us that most of those capacities (in fact, all of them save one) are essentially embodied. Each involves some bodily state with a form:

> In fact, the affections of the soul all seem to be with a body: excitement, calmness, fear, pity, courage, also joy, and both loving and hating, for whenever we have them the body is in some way affected . . . If this is the case, then clearly the affections are enmattered forms (*logoi enhuloi*). Hence, their definitions will be for instance, "anger is a certain movement of such-and-such a body (or a part or power of it) by this for the sake of that" . . . [T]he affections of soul are inseparable from the natural matter of animals. (403a16–19, 24–7; 403b17–18)[3]

On Aristotle's view, states like anger are physiological processes with distinctive forms or structures. As Richard Sorabji says, "[anger's] material cause is specified as a physiological process, but [its] formal cause relates it to another capacity: desire. For the formal cause of anger is a desire to retaliate" (1992: 208).

Against the background of his hylomorphic theory, Aristotle's claim that *nous* is not essentially embodied appears anomalous, and many commentators have found the anomaly quite puzzling. Marjorie Grene is an example:

> Aristotle did believe that rational mind . . . was different, in its independence of body, from any other aspect of 'soul'. And I can only record, further, my own bafflement by this view. Platonic separate soul, Platonic immortality one must take seriously as Plato's deepest faith . . . [But n]o one could reject more emphatically than Aristotle a mind- (or soul-) body dualism in general . . . Soul is the way the body works: that is all. When it comes to *νοῦς*, however, the case is different . . . This seems a strange deviation from the principal tenor of his own psychology. Everywhere else we have soul-body unity . . . it is the soul-in-the-body, soul as form *of* body, that is concerned in all cognitive activities except for the purely rational and immediate activity of *νοῦς*: in all activities—and he knew they were many—that we share with other animals. Then suddenly with our power . . . to grasp first principles . . . comes another element in our natures,

[3] He reiterates the point after developing his account of soul in Book II: "It is clear, then, that the soul (or some parts of it if it is divisible) cannot be separated from the body, for in some cases the actuality of the [soul's] parts is the actuality of [the body's] parts themselves" (413a4–6).

> foreign to body altogether . . . I cannot help feeling it strange that we should do this so suddenly and briefly at the very top of our biological ladder. It seems to me much more plausible to suppose that far down in the history of life itself something like knowing began: that our reaching out to reality is but a fuller development, marvelously magnified by the invention of symbolic systems, of a curiosity common to many kinds of animals, supported by different sensory capacities, different kinds of awareness, different passions, but fundamentally akin. (1963: 243–5)[4]

The embodiment thesis thus represents the default hylomorphic position, even for hylomorphists who reject it, like Aristotle. This observation has an important dialectical implication: if the essential embodiment of our capacities is the default hylomorphic position, then it is not the truth of that thesis which requires special defense, it is rather its falsity which does. If some of our capacities are not essentially embodied, then the burden is on embodiment's detractors to establish this.

Most of the arguments that have been advanced against the embodiment thesis are descendants of the argument from *De Anima* III.4 quoted above.[5] They focus on thought or understanding, and claim that our capacity to think or understand is unlimited in its scope in a way that an embodied capacity cannot be. We are capable of thinking about anything, it seems, but we are not capable of perceiving just anything. Our eyes are sensitive only to light, not to sound, our ears are sensitive only to sounds, not to flavors, and even then, they are sensitive only to sounds within a limited range of frequencies. These kinds of limitations on human perceptual capacities do not appear to be mirrored in the case of human thought or understanding. We may not be able to perceive x-rays or hear sounds with frequencies above 22 kilohertz, but we can still think about x-rays and understand that there are sounds with frequencies above 22 kilohertz. What explains this difference between perception and understanding? It seems evident that the limitations on our perceptual capacities are imposed by our sensory organs. This is certainly Aristotle's own view. According to him, each organ is an intermediate state (*mesotēs*), a mean between extremes, which enables it to be sensitive to the characteristics of things in the environment, or as he puts it, which enables it to receive the sensible forms of things without their matter (*aneu tēs hulēs*) in something analogous to the way a

[4] Other commentators who find Aristotle's denial of essential embodiment puzzling include Werner Jaeger (1934: 332–4), W. D. Ross (1957: 65–7), Edwin Hartman (1977: 252), Kathleen Wilkes (1978: 115–16), and John Sisko (1999: 264–6).

[5] This is certainly true of the arguments advanced by medieval Aristotelians like Aquinas. Here is his formulation of the argument: "What can know things must have nothing in its own nature because what is in it naturally would impede the knowledge of other things. Thus we see that a sick man's tongue, which is infected with a feverish and bitter humor, is not able to perceive something sweet, but everything seems bitter to it. Every body, however, has some determinate nature. It is impossible, therefore, that the intellectual principle should be a body, and similarly impossible that one should understand through a bodily organ, for the determinate nature of the bodily organ would prohibit knowledge of every body, just as if some determinate color is not only in the pupil of the eye but also in a glass vase, the liquid in the vase seems to be that color" (*Summa Theologiae* Ia, Question 75, Article 2 *sed contra*).

piece of wax receives the shape of a ring irrespective of what the ring is made of (424a19–24).

Understanding is like perception, Aristotle thinks, insofar as both involve receiving the forms of things without their matter. I can grasp what it is to be a triangle without thinking specifically of a gold, or wood, or bronze triangle. But understanding is unlike perception in that its scope is not similarly limited. Not only can we understand things that we cannot see or hear, but we can grasp what a triangle is independent of the imperfections of this or that particular triangular object. This difference, Aristotle thinks, shows us that understanding is not embodied, that it has no organ, and is in fact entirely unmixed with a body.

Aristotle advances at least two arguments in favor of this conclusion. First, he tells us that the intrusion of anything foreign to understanding would hinder and obstruct (*kōluei kai antiphrattei*) its operation. The reasoning behind this claim is based once again on the analogy with perception. On Aristotle's account, a sensory organ is an intermediate state between sensory extremes, and it is by altering this state—moving it toward one extreme or the other—that a perceptible object is able to affect the sense organ and thereby actualize our power to perceive. The organ of touch, for instance, is in an intermediate state between extremes of hot and cold, and because of that, it is able to be affected by things that are both hotter and colder than itself, for objects of either sort will move the organ's state toward one or the other extreme. For the same reason, Aristotle says that the eye-jelly (*korē*) must be transparent if we are to be capable of seeing the range of colors we do (*De Sensu* 438a13–17): the transparent is colorless, and what is colorless is capable of receiving any color (*De Anima* 418b26). This account implies that there are limitations on perceptual capacities. Plants are altogether incapable of perception because they have no parts in the relevant intermediate states (424a32–b2), and can thus be affected by forms only with the matter (*meta tēs hulēs*). They can become hotter or colder, for instance, but they cannot perceive hot or cold. Moreover, since we can only perceive things that are capable of moving the intermediate states of our sensory organs toward some extreme or other, we cannot perceive an object that is just as hot or cold as we are: such an object is incapable of moving the organ's intermediate state toward either extreme (424a2–5).

Suppose now that understanding were embodied like perception. If that were the case, Aristotle reasons, then the organ in which it was embodied would have to work in something like the way an organ of perception does. It would have to occupy an intermediate state that was capable of being altered by the forms of objects in the environment—not sensible forms, however, but intelligible ones. Yet, if that were the case, it seems that our capacity for understanding would have restrictions in just the way our capacity for perceiving does. Just as the organ of touch can't be affected by anything that is as hot or cold as it is, a circumstance that renders us incapable of perceiving such things, likewise an organ of understanding could not be affected by anything that had the same kind of intermediate state as the organ itself, for that

would render us incapable of understanding such things. But in fact, says Aristotle, we are not incapable of understanding anything bodily. Our capacity for understanding is not limited the way perception is. Consequently, understanding must not be embodied the way perception is.

Aristotle's second argument appeals to the destructibility of the senses:

> after a sense perceives something very intense, it cannot perceive; for instance, after hearing very loud sounds it cannot hear sound, and after seeing vivid colors or smelling strong odors, it cannot see or smell. But whenever understanding grasps something that is very intelligible it grasps more, not less, about inferior things. (429b1–5)

Our perceptual capacities can be destroyed by very strong sensations, says Aristotle; the reason is that they are embodied: very strong sensations destroy the intermediate state of a sensory organ, just as plucking the strings of the instrument too forcefully destroys its tuning (424a30–2). If understanding were embodied like perception, then it too could be destroyed by excess, but understanding is not destroyed by excess. Hence, Aristotle concludes, understanding is not embodied like perception.

These, then, are Aristotle's arguments against the embodiment thesis. What can defenders of embodiment say in response? First, it's important to remind ourselves of the anomalous character of *nous* within Aristotle's hylomorphic framework. The problem is not merely that *nous* stands strangely apart from the rest of our capacities on Aristotle's account; it is, rather, that it's not clear that Aristotle's account of *nous* fits within his general hylomorphic framework at all. Christopher Shields states the worry as follows:

> In its original context, hylomorphism held that every change involved a complex, something continuing, for example the bronze underlying the generation of a *statue*, and something gained or lost, for example a *Hermes-shape*. Now we learn that the mind does not exist before it thinks. It seems to follow that this is not amenable to hylomorphic analysis . . . [H]ylomorphism requires the existence of a persisting subject, and the mind cannot persist through a change if it does not exist in actuality before the change begins [contrary to what Aristotle says at 429a22] . . . A defender of Aristotle might . . . reasonably point out that . . . it exists in potentiality. That is fair enough, but in one sense that only postpones the question. It seems fair to say that *x* can exist only in potentiality, but only if there is some *y* which exists in actuality, and *x* is a potentiality grounded in *y* . . . If that is so, however, then Aristotle owes an account of what precisely exists in actuality such that it has the capacity which is made something actual only when thinking begins . . . Aristotle has some room to respond here, but it is fair to say that he has now taxed his hylomorphic theory of change almost beyond the limits of recognition. (2014: 354–5)

It is not altogether clear, therefore, that a hylomorphic framework has the kind of plasticity that critics of essential embodiment need it to have. A theory which allows for disembodied human capacities is in danger of being hylomorphic only in name.

Let us turn now from the tension in Aristotle's overall framework to the argument of *De Anima* III.4 itself. Here things don't seem to be much better for Aristotle.

Charles Kahn remarks that "Aristotle's own arguments here are surprisingly weak and insubstantial, as if surrounded by Platonists rather than materialists, he did not regard this position as controversial enough to stand in need of a real defense" (1992: 375–6). To appreciate this, let us assume that understanding is like perception in the way Aristotle describes: it involves receiving the intelligible forms of things without the matter in the way that perception involves receiving the sensible forms of things without the matter. In that case, it still doesn't follow that an organ of understanding would have to work in the way an organ of perception does. This seems especially true given the differences between intelligible and sensible forms on Aristotle's account. What is intelligible about a thing, Aristotle says, is its essence, what-it-is-to-be (*to ti ēn einai*) that kind of thing. I understand what a human is by grasping what it is to be a human. This is something universal, not particular. By contrast, when I perceive this or that human, what I perceive is particular: I see or hear these or those particular colors or sounds. It is because perception deals with particulars that the account of the sense organ occupying an intermediate state has any plausibility. That state makes it possible to explain how particular colors or sounds trigger the activation of the sense organ and thus enable us to perceive. By contrast, it does not seem very plausible at all to suppose that something universal, such as the essence of humanity, should trigger an organ in an analogous way. Defenders of Aristotle's argument might insist that this is precisely the point, that the difference between intelligible things and sensible ones is precisely what supports the conclusion that understanding isn't embodied the way perception is. But it is possible to draw a different lesson: the basic analogy between understanding and perception is ill-founded. It's true that both understanding and perception involve openness to the surrounding world in a way that other interactions with the world do not. Becoming hot or cold is different from becoming aware of hot or cold, or understanding what hotness and coldness are. But this similarity seems insufficient to ground the claim that if there are mechanisms of understanding, their operations must mirror the operations of perceptual mechanisms.

There is a second way of criticizing Aristotle's argument. Suppose that we take the analogy between understanding and perception seriously. Both, we acknowledge, open us up to the world in a special way. In the case of perception, that openness is achieved through the operation of physiological mechanisms. If understanding is analogous to perception, then that would seem to indicate that the openness of understanding can be achieved through the operation of physiological mechanisms as well. As we've seen, Aristotle thinks that the distinctive character of intelligible forms rules this out, but if the intelligible forms are unlike sensible forms in the way Aristotle supposes, then the analogy between understanding and perception is not really very deep, certainly not deep enough to ground the claim that mechanisms of understanding, if they exist, must operate in the way mechanisms of perception do.

A third difficulty with Aristotle's argument concerns the two ways in which something can receive form: with the matter versus without it. Consider again

Aristotle's argument that the eye-jelly must be transparent: if it were a determinate color such as red, then that color would prevent the eye from receiving the form of redness. The eye cannot receive a form it already has, so red eye-jelly would prevent us from seeing red things. The key assumption in this argument is that having a form with the matter prevents that thing from receiving the form without the matter. Having redness with the matter, that is *being* red, prevents something from receiving redness without the matter, that is *perceiving* red. It is plausible to suppose that something that already has a form in either sense cannot receive the form in that same sense. If something already is red, then it cannot become red, and if something already perceives red, then it cannot come to perceive red. But why suppose that something that already is red cannot come to perceive red if the kind of reception of form involved in perception is different from the kind of reception of form involved in literally being red?

Richard Sorabji (1974, 1992) has interpreted Aristotle as claiming that the eye-jelly literally turns red when we perceive a red object. His interpretation can be criticized on several grounds, but this much can be said in its favor: it makes it clear why being red would interfere with perceiving red, for on Sorabji's interpretation, perceiving red involves literally becoming red, and something cannot literally become red if it already literally is red. If, by contrast, a literalist interpretation like Sorajbi's is incorrect, then it remains unclear why we should endorse the assumption on which Aristotle's argument is based, why we should believe that having form F with the matter prevents something from receiving F without the matter. What is true of perception is also true of understanding. If nothing prevents an organ of perception from having F with the matter and yet receiving F without the matter, then nothing should prevent an organ of understanding from doing the same. In that case, though, Aristotle's argument collapses, for an organ of understanding could have a number of forms with the matter while yet having no forms without the matter prior to performing an actual act of understanding.

A final, related objection challenges the basic account of perception and cognition on which Aristotle's argument is based. That account is based on the idea that the perceptive or cognitive faculty takes on the form of something without the matter. For Aristotle, taking on form either with the matter or without it involves instantiating, exemplifying, or otherwise having that form. But is this really a plausible way of understanding perception or thought? Fred Miller articulates the worry as follows:

> Crucial for Aristotle's argument is the assumption that thinking, like perceiving, is receiving a form . . . However, there are two different ways of receiving the form of cold. One is receiving the cold literally or physically and becoming cold. The other way is receiving it 'objectively', that is, having had it become the object of one's awareness. The latter is clearly what Aristotle means . . . Thinking of an object is like perceiving an object in that both are ways of representing an object . . . One way [of representing an object] is to simply copy or imitate the object . . . But it is not necessary to exemplify a characteristic directly or literally in order to represent it. Encoding a characteristic is a way of representing an object without exemplifying it[;] for

example, the formula H_2O encodes water without exemplifying water or even exemplifying the structure of water. If it were assumed that when it perceives an object the perceptive faculty in some way exemplifies it . . . and it is assumed that thinking resembles perceiving in this way, then Aristotle's . . . premise seems plausible. These assumptions are, however, extremely controversial. Why suppose that the concept of water exemplifies aqueousness in even an attenuated way, any more than the formula H_2O exemplifies aqueousness? (2012: 326–7)

Let the foregoing criticisms suffice for a discussion of Aristotle's first argument. Consider now his second: (1) If understanding is embodied, then it can be destroyed by excess, but (2) understanding cannot be destroyed by excess; therefore, understanding is not an embodied capacity. We've seen that the case for Premise (1) is based on the following line of reasoning: the senses can be destroyed by excess (loud sounds are deafening, bright lights are blinding) because they are embodied. Understanding is like perception; hence, it too can be destroyed by excess if it is embodied. Since this line of reasoning is based on the analogy between perception and understanding, it is subject to the criticism mentioned earlier: given the different natures of intelligible and sensible forms, it is not clear why an organ of understanding would have to work in the way an organ of perception does.

We've seen, moreover, that the argument for Premise (2) is based on the idea that intelligible objects will affect us in something analogous to the way perceptible objects do—in particular, it assumes that we can speak of something being highly intelligible in the way we can speak of a sound being very loud or a light being very bright, but it's not clear that this is really the case. D. W. Hamlyn remarks that what Aristotle means in speaking of an object as very intelligible (*sphodra noēton*) here is "most obscure" (1968: 137). The clearest indication of what Aristotle has in mind is his claim that something very intelligible enables us to grasp more about inferior things. Elsewhere, Aristotle says something similar about true principles: he tells us that a true principle enables us to see particular matters of fact in proper perspective (*Nicomachean Ethics* 1098b12–13). But if this is the kind of thing Aristotle has in mind, it is difficult to sustain the analogy between understanding and perception for at least two reasons. First, the way we come to grasp a theoretical principle is completely unlike the way we come to perceive a bright light or loud sound. The latter processes are passive in a way that the former is not. Second, it is unclear that our ability to understand things isn't destroyed by excess. Sometimes when learning something new, we reach a point of information saturation: we cannot absorb anything further and we need to rest. In addition, if by speaking of very intelligible objects Aristotle means something like true principles, then it becomes unclear that our ability to understand things is really unlimited in the way his main argument assumes. Perhaps some of us have inbuilt cognitive limitations that prevent us from ever grasping certain first principles. In fact, Aristotle suggests that there are cases of precisely this sort. People who have been raised with poor moral habits, he tells us, make unsuitable moral interlocutors because they are not able to grasp the right moral principles (1095a32–b13).

These are just some of things that defenders of the embodiment thesis can say in response to Aristotle's argument. I conclude that the argument does not succeed in disproving the thesis: it does not give us sufficient grounds for rejecting the default hylomorphic position that all our capacities are essentially embodied.

Is there a positive argument to be made for the embodiment thesis? I think there is. It's suggested by Marjorie Grene's remarks: the idea that all of our capacities, including our capacity to think, are essentially embodied is more plausible than its denial, given what we know about human evolution. Think again of our capacities to grow and develop, to crawl, walk, and talk, to reach, grasp, run, and jump, to throw, catch, taste, swim, see, hear, eat, breathe, and so on. In each case, the human capacity evolved from ancestral forms in conjunction with cultural forces. Against this background, it is plausible to conclude that in all likelihood the capacity for thought is no different, that the cultural development of symbolic systems gave rise to forms of social and environmental interaction which created conditions that selectively favored individuals who were more adept at modulating their behavior through the use of those systems. To speak of having a mind, then, is, as Dewey says, to speak of "an added property assumed by a feeling creature, when it reaches that organized interaction with other living creatures which is language, communication" (Dewey 1958: 258).

Even in a nonevolutionary framework like Aristotle's it is possible to run an argument that essential embodiment is more plausible than its denial, given that so many of our capacities seem so obviously embodied. If our powers to grow, crawl, walk, talk, see, hear, and so on are all essentially embodied, then it seems plausible to generalize that all our powers are essentially embodied. Naturally, because the argument is inductive, it's possible to argue against the conclusion. We've seen, however, that the traditional Aristotelian argument against the embodiment thesis falls short. There might be other arguments, but given the vast range of our embodied capacities, and given that the embodiment thesis represents the default hylomorphic position, the burden of proof rests squarely on the shoulders of embodiment's opponents.

8.4 Psychological Activities

If all our capacities are essentially embodied, hylomorphism implies that thinking, feeling, perceiving, intentionally acting, and other activities that philosophers typically categorize as mental or psychological are all structured manifestations of powers. When I think, feel, perceive, or act, I coordinate the manifestation of the powers of my parts and in some cases the powers of surrounding materials as well. The coordinated manifestation of those powers composes my thoughts, feelings, perceptions, and actions in the way described in Section 8.2. An example will help illustrate this idea.

Perceptual states are often categorized as mental states. According to hylomorphists, when Gabriel sees something—the lock on a door, say (or more precisely,

physical materials that are spatially arranged lock-wise)—he and the lock both manifest powers they possess. He manifests the power to see the lock, and it manifests the power to be seen by him. Gabriel and the lock are reciprocal disposition partners in Martin's (1997, 2007) sense (Section 4.1). The powers of both are mutually manifested in each other's presence when the surrounding conditions are right, just as water and salt mutually manifest their powers to dissolve and be dissolved, respectively, when conditions are right.

Gabriel has the power to see by virtue of having the parts he has (Sections 8.2–8.3). The coordinated manifestations of the powers of some of his parts (intuitively, those composing his visual system) contribute to his seeing, and those parts form a subset of the individuals whose powers, when manifested in the right way, compose his seeing. The same is true mutatis mutandis of the lock. Intuitively, the parts of the lock, by virtue of which it has the power to be seen, are those composing its surface, the ones which reflect light to Gabriel's eyes.[6] In addition, there are other environmental factors involved in Gabriel's seeing the lock, such as the direction and intensity of the light, the condition of the air through which he sees it, and so on. Gabriel's seeing the lock is thus a complex structured activity composed of the coordinated manifestation of the powers of his parts and those of the surrounding materials. Those manifestations are in turn powers for further manifestations—the power, for instance, to lock the door with a minimum of fumbling, or quickly to navigate the hallway and stairs.

The foregoing example illustrates the way hylomorphists look to understand perceptual states in terms of the structured manifestation of powers. Consider another example, one that illustrates a different kind of psychological structure. Suppose that I am helping Gabriel move some items out of his second-floor apartment. Because he lives in a questionable neighborhood, his door is well outfitted with a number of locks. It's a short trip down the stairs, and we intend on coming back as soon as we've carried one of the items out to his car, which is parked directly in front of the building. Before proceeding down the stairs, however, he stops, removes a large ring of keys from his pocket, and proceeds to lock the door. I find this behavior puzzling. "Why are you locking the door?" I ask, "We're coming right back up." He responds: "I'm locking the door because I was recently robbed while on a short trip down the stairs just like this one. And," he adds *sotto voce*, "I suspect my neighbor across the hall may have been the culprit." His answer resolves my puzzlement. Why? Intuitively, the reason is that it enables me to discern a rational structure in his behavior that I could not discern before. We get a sense for what that structure is by considering the assumptions about Gabriel and his circumstances that generate the initial puzzle:

(1) Gabriel is a rational being, one whose behavior can be explained by appeal to reasons.

[6] More precisely, they are the materials such that if there were a lock, they would compose its surface.

(2) Behavior that can be explained by appeal to reasons tends to approximate rational coherence.
(3) Gabriel wants to save time and unnecessary effort.
(4) Locking the door in the present circumstances needlessly wastes time and effort.
(5) Gabriel is locking the door.

Together claims (1)–(5) generate a tension that I do not initially know how to resolve. If Gabriel is a rational being who wants to save time and effort, and locking the door is a waste of time and effort, then it seems he shouldn't be locking the door. Because claims (1)–(5) all seem to be true, I am at a loss how to make sense of Gabriel's behavior. Gabriel's answer provides the key. It challenges my commitment to claim (4). It suggests that locking the door is not a needless waste of time and effort since Gabriel does not want to be robbed, and the chances of being robbed in the time it will take for us to return are fatter than I initially assumed. I thus recognize a rational structure in Gabriel's behavior which I was previously unable to discern, a structure describable in the following terms:

> Gabriel is locking the door because he's afraid he might get robbed by his neighbor—something he believes likely on the basis of recent events.

We can easily imagine other answers to my question that would have disclosed different rational structures in Gabriel's behavior. Here are some examples:

(a) "I've lately been trying to cultivate an attitude that is heedless of the constraints of time and effort";
(b) "I'm just checking that this key turns in the lock. The outside door of the building is keyed the same way, and this is not the key I usually use";
(c) "I really just enjoy locking doors; I love the feel of the lock snapping into place and the satisfying sound it makes when it does so";
(d) "Actually, before returning we have to go down the block and pick up something from the corner store";
(e) "Physicists have discovered a temporal anomaly in this building: people walking up the stairs take three times longer to cover the same distance as people outside the anomaly";
(f) "Whoops! Sorry, there is no reason, just force of habit."

Any of these answers could resolve my puzzlement. Each implies the falsehood of one of my initial assumptions, and so each (with the exception of (f), which reveals that his behavior wasn't entirely rational) enables me to discern a rational structure in Gabriel's behavior—structures we might describe in the following ways:

- Gabriel is locking the door as part of a spiritual exercise.
- Gabriel is not locking the door, he is merely testing the key.
- Gabriel is locking the door for the sheer thrill of doing so.

- Gabriel is locking the door because he's afraid he might get robbed by his neighbor—something he believes likely on the basis of recent events and the amount of time we'll be gone (which is longer than I initially assumed).
- Gabriel is locking the door because he's afraid he might get robbed by his neighbor—something he believes likely on the basis of recent events and what physicists have recently discovered about his building.
- Gabriel wasn't acting entirely rationally: he wasn't thinking about what he was doing, but was just locking the door out of habit.

These examples illustrate other ways in which human activity can be structured, ways that are subject to rational evaluation. In each case, what gets structured in these ways is a range of Gabriel's activities: his seeing, his grasping the key, his turning it in the lock, and so on.

There are several noteworthy implications of understanding thought, feeling, perception, and intentional action in terms of a hylomorphic framework of powers and manifestations. One was noted earlier in connection with other structured activities: lower-level behaviors needn't compose higher-level activities. Just as my muscular contractions and neural firings needn't compose an act of throwing a baseball, likewise the activities of Gabriel's parts needn't compose a perceptual activity, or a rational activity, or an activity with some other psychological structure. The case in which Gabriel acts merely out of habit is an example. His locking the door fails to satisfy certain criteria for rational evaluation. Consider likewise hallucinations. A mad neuroscientist might hotwire Gabriel's visual system and coordinate the activation of its parts in a way that exactly mimics their activation when Gabriel sees a lock. But in the absence of a real lock, the parts of Gabriel's visual system would not manifest their powers in a way that composed a perceptual activity. Gabriel would not be perceiving a lock, he would be having a hallucination as if he were seeing a lock.

The foregoing sections have indicated how activity-making structures matter, how they make a difference, and how they count. The discussion in this section indicates in a preliminary way how activity-making structures *mind*, how they provide us with resources for accommodating thought, feeling, perception, and intentional action within the natural world.

8.5 The Mental-Physical Distinction

One of the more striking implications of the hylomorphic approach to thought, feeling, perception, and intentional action is that it does not treat the mental-physical distinction as canonical, but simply as an artifact of our descriptive and explanatory interests. Mind-body problems are often characterized as problems understanding how mental phenomena are related to physical phenomena. Hylomorphists, however, look to understand thought, feeling, perception, and intentional action—and our

activities and capacities generally—in terms of the structured manifestations of powers. In a hylomorphic framework, whether we decide to call some of these powers or manifestations 'mental' or 'nonmental,' or 'physical' or 'nonphysical', is orthogonal to the project of understanding what they are, and why and how they operate as they do.

There are many ways of drawing a distinction between mental and nonmental, and physical and nonphysical. Here are just a few:

- Something is mental if and only if it displays intentionality.
- Something is mental if and only if it is subjective or has a subjective point of view.
- Something is mental if and only if it can be described or explained using a psychological vocabulary.
- Something is mental if and only if we have privileged access to it.
- Something is physical if and only if it can be exhaustively described by physics.
- Something is physical if and only if it is composed of materials that can be exhaustively described by physics.
- Something is physical if and only if it belongs to the causal order of the world.
- Something is physical if and only if we do not have privileged access to it.

Hylomorphists are free to adopt definitions like the foregoing as suits their descriptive and explanatory purposes, but nothing about their framework of powers and manifestations forces them to adopt one of these definitions as opposed to another, or to adopt any such definitions at all.

I have argued elsewhere, and I will argue again later, that this feature of hylomorphism enables hylomorphists to solve some traditional mind-body problems in very attractive ways (Jaworski 2011, 2012). Other philosophers have had the idea that an adequate resolution of mind-body problems will require a radical revision of our concepts, perhaps even a revision that rejects the mental-physical dichotomy. A recent example is John Searle:[7]

> [Conceptual dualism] consists in taking the dualistic concepts very seriously . . . Both traditional dualism and materialism presuppose conceptual dualism . . . [M]aterialism inadvertently accepts the categories and the vocabulary of dualism. It accepts the terms in which Descartes set the debate. It accepts . . . the idea that the vocabulary of the mental and the physical, of material and immaterial, of mind and body, is perfectly adequate as it stands . . . What I believe . . . is that the vocabulary, and the accompanying categories, are the source of our deepest philosophical difficulties. As long as we use words like "materialism," we are almost invariably forced to suppose that they imply something inconsistent with naïve mentalism. I have been urging that . . . [o]ne can be a "thoroughgoing materialist" and not in any way deny the existence of . . . mental phenomena. However, since my use of these terms

[7] Another example is John Dewey: "[T]he 'solution' of the problem of mind-body is to be found in a revision of the preliminary assumptions . . . which generate the problem" (1958: 263).

runs dead counter to over three hundred years of philosophical tradition, it would probably be better to abandon this vocabulary altogether. (1992: 26, 54–5)

Searle suggests that resolving mind-body problems requires abandoning the traditional philosophical vocabulary that distinguishes the mental and the physical. When it comes to describing his own view and others', however, Searle does not abandon any of the traditional categories. The idea that he takes to be innovative, namely that someone can be a thoroughgoing materialist and yet not deny the existence of the mental, is the central claim of the psychophysical identity theory, arguably the paradigmatic materialist view. Psychophysical identity theorists claim that mental phenomena are identical to physical phenomena; they are thus thoroughgoing materialists, yet they in no way deny that mental phenomena exist. Searle says, moreover, that according to his own theory, mental states are both caused by and realized in physical states of the brain (2004: 79). To articulate his own theory, therefore, he uses the very same dichotomous mental and physical categories that he claims are responsible for generating mind-body problems.

There might be many reasons why Searle doesn't abandon the traditional mind-body categories when articulating his theory, but perhaps the most obvious is that he doesn't develop a metaphysical framework that would enable him to formulate his theory in a different way. He doesn't dig deeply enough into the metaphysical soil to uproot the mental-physical dichotomy that he and other philosophers take to be problematic. Hylomorphism, by contrast, does. The hylomorphic framework does not imply a commitment to categorizing powers or manifestations, or the systems or subsystems that have them, as mental or nonmental, physical or nonphysical. According to hylomorphists, we can get on with our empirical investigation of powers, parts, and manifestations without ever employing those categories.

Consider again an activity like Gabriel's running, or his seeing the lock, or his turning the key in it. We've seen that according to hylomorphists, activities like these comprise many different contributing factors—both factors within the organism and factors in the surrounding environment. According to hylomorphists, we select the factors that interest us, ignore others, and proceed to investigate the ones that interest us empirically by the appropriate methods. By analogy, it typically takes a whole range of different factors to cause a car crash: the balding tires, the faulty brake mechanism, the inadequate roadway grading, the driver's blood-alcohol level—all of these and more contribute to the crash; all of them belong to what J. S. Mill (1843: Book 3, chapter 5, section 3: 398ff.) would call 'the complete cause' of the crash. In most contexts, however, we are not concerned with describing something's complete cause, but only a handful of contributing factors. We focus on those and ignore the rest. David Lewis describes the idea this way:

We sometimes single out one among all the causes of some event and call it "the" cause, as if there were no others. Or we single out a few as the "causes," calling the rest mere "causal factors" or "causal conditions." Or we speak of the "decisive" or "real" or "principal" cause. We

select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we want to talk about. (1973b: 162)[8]

The automotive engineer focuses on the brake mechanism, the civil engineer on the roadway grading, the prosecutor on the blood-alcohol level, and so on.

According to hylomorphists, something analogous is true of human behavior: we select the factors that interest us, investigate those, and ignore the rest. Eleanor's running involves the coordinated manifestation of the powers of various physiological systems and subsystems. Among them, we might choose to focus on her circulatory system, which contributes to her running by supplying oxygenated blood to her muscles. Functional analysis reveals that the activity of that system involves the coordinated manifestation of the powers of various further subsystems. Among these, we might choose to focus on the heart. Analysis of its pumping activity shows that it is composed of muscle tissues that undergo coordinated contraction and relaxation. Analysis of the contractions and relaxations might reveal contributing subactivities at a lower level still, and we might choose to focus on this or that contributing subactivity of this or that lower-level subsystem. But perhaps we are not interested in lower-level subsystems at all. Perhaps we are interested in the more inclusive rational activities into which Eleanor's running is integrated: Eleanor is running because she wants to stay fit, and she wants to stay fit because she wants to continue playing softball at a competitive level, and she wants to continue playing softball at a competitive level because...

Real human behavior comprises so many structural levels that a complete account of what we would typically consider a simple intentional action would have to describe a complex range of social, environmental, and physiological factors. In most contexts, however, we are not concerned with describing the complete cause of someone's behavior; we instead have a limited set of descriptive and explanatory interests in mind, and focus on a small subset of contributing factors. We'll see in Chapter 13 that this view of explanation has important implications for the way hylomorphists approach the problem of mental causation. For the time being, however, the important point is that the mental-physical distinction is foreign to the basic categories in terms of which hylomorphists understand human behavior. At no point in the empirical investigation of human activities and the factors that

[8] Elsewhere, Lewis expresses the idea as follows: "The multiplicity of causes and the complexity of causal histories are obscured when we speak, as we sometimes do, of *the* cause of something. That suggests that there is only one... If someone says that the bald tyre was the cause of the crash, another says that the driver's drunkenness was the cause, and still another says that the cause was the bad upbringing that made him reckless, I do not think that any of them disagree with me when I say that the causal history includes all three. They disagree only about which part of the causal history is most salient for the purposes of some particular enquiry... Some parts will be salient in some contexts, others in others. Some will not be salient in any likely context, but they belong to the causal history all the same: the availability of petrol, the birth of the driver's paternal grandmother, the building of the fatal road, the position and velocity of the car a split second before the impact" (Lewis 1986c: 215–16).

contribute to them are we compelled to introduce a distinction between mental factors and nonmental ones, or between physical factors and nonphysical ones. Nothing forces hylomorphists to accept a mental-physical distinction. Nothing forces them to reject one either. The choice to accept or reject such a distinction is contingent upon whatever theoretical or practical ends hylomorphists are looking to achieve. Within a hylomorphic framework, the mental-physical distinction is an artifact of our descriptive and explanatory interests.

8.6 Conclusion

Individual-making structures are not the only structures that exist according to hylomorphists. The activities in which structured individuals engage have structures of their own—activity-making structures. These have many of the same characteristics as individual-making structures: they matter, they count, they make a difference, and they also mind. Hylomorphists take thought, feeling, perception, and intentional action to be species of structured activities. They are composed of the structured manifestations of the powers of an individual's parts and surrounding materials. The powers of a structured individual are thus embodied in its parts: it has its powers by virtue of having the parts it has since the structured manifestations of the powers of those parts (in conjunction perhaps with the powers of surrounding materials) compose the activities of the structured individual as a whole.

Are all of an individual's powers essentially embodied? The embodiment thesis claims that all of them are. Aristotle rejected the embodiment thesis, and many hylomorphists have followed his lead. The argument in *De Anima* III.4 purports to show that our capacity to think or understand is not essentially embodied. That argument, however, is far from decisive, and defenders of embodiment have several ways of responding. In addition, there is a positive argument in favor of the embodiment thesis: the essential embodiment of all our capacities fits much more comfortably within the overall picture of human beings as products of natural selection. Even within a nonevolutionary framework, the essential embodiment of all our capacities receives strong inductive support, given the range of human capacities that seem so obviously embodied.

One of the more striking implications of the hylomorphic approach to thought, feeling, perception, and intentional action is that the mental-physical distinction is not treated as canonical, but simply as an artifact of our descriptive and explanatory interests. Mind-body problems are often characterized as problems understanding how mental phenomena are related to physical phenomena. Hylomorphists, however, look to understand thought, feeling, perception, and intentional action in terms of the structured manifestations of powers. Whether we decide to call some of these powers or manifestations 'mental' or 'nonmental,' 'physical' or 'nonphysical' is orthogonal to understanding what they are, and why and how they operate as they do.

9

Hylomorphic Necessitation and Supervenience

9.1 Activity and Power Necessitation

Recall that according to the embodiment thesis defended in Section 8.3, all the powers of a structured individual are essentially embodied in its parts. A hylomorphic theory committed to the embodiment thesis implies some robust supervenience and necessitation theses. Supervenience is a kind of property covariation relation. To say that A-properties supervene on B-properties is to say that things cannot differ in their A-properties without also differing in their B-properties, or equivalently, B-twins must be A-twins. Some aesthetic properties, for instance, supervene on physical properties: two individuals cannot differ from each other aesthetically without differing from each other physically. If painting a is physically indistinguishable from painting b, if they have exactly the same physical features, then it is impossible for one to be beautiful and the other ugly, or for one to be well proportioned and the other poorly proportioned. If aesthetic properties supervene on physical properties, then necessarily, physical twins must be aesthetic twins: any aesthetic differences between a and b must be traceable to physical differences between them.[1]

There are stronger and weaker supervenience theses. Here are two examples:

Weak supervenience: For any possible world w, and any x and y in w, if x and y are B-indistinguishable, then x and y are A-indistinguishable.

Strong supervenience: For any possible worlds w_1 and w_2, and any x in world w_1 and y in world w_2, if x and y are B-indistinguishable, then x and y are A-indistinguishable.

The difference between these supervenience theses is that weak supervenience specifies a covariation relation only within a single world. In any one world, it says,

[1] Importantly, this does not imply that aesthetic properties are explained by physical properties. If aesthetic properties are correlated with physical properties in the way supervenience says, it does not necessarily follow that something has the aesthetic properties it has *because* it has the physical properties it has. The idea that supervenience implies an explanatory condition is one of the most common mistakes that has been made in discussions of supervenience. I will have occasion to discuss this mistake in greater detail in Section 10.1.

B-indiscernibility guarantees *A*-indiscernibility. But knowing that *A* and *B* properties covary in one world tells us nothing about whether and/or how they covary in others. As a result, weak supervenience is compatible with the possibility that a *B*-replica of something might have radically different *A*-properties or perhaps no *A*-properties at all. Strong supervenience, by contrast, rules this out. Because *x* and *y* can belong to different possible worlds, it specifies a covariation relation that obtains across worlds, not just within a single world. It says that for any individuals in any worlds at all—individual *x* in world$_1$ and individual *y* in world$_2$—*B*-indiscernibility guarantees *A*-indiscernibility. *B*-twins, in other words, will be *A*-twins, even if they are in different worlds. Strong supervenience thus implies weak supervenience, but not vice versa.

Necessitation is a kind of property covariation relation as well, but it is typically stronger than even strong supervenience:

> *Necessitation*: Necessarily, for any *x*, if *x* has an *A*-property, *F*, then there is a *B*-property, *G*, such that *x* has *G*, and necessarily, for any *z*, if *z* has *G*, then *z* has *F*.

At one time, the difference between supervenience and necessitation was not clearly appreciated. Jaegwon Kim (1987, 1990), for instance, claimed that necessitation and strong supervenience were equivalent and used 'strong supervenience' as a label for both. It is true that necessitation implies strong supervenience, provided that necessity is understood as quantification over possible worlds,[2] but there is good reason to think that the converse does not obtain: strong supervenience does not imply necessitation. Kim's claim that it did was based on various assumptions about properties, including the assumption that properties are closed under the Boolean operations of conjunction, disjunction, and complementation. But a theory of properties is not beholden to that assumption, and in fact, the theory of properties

[2] To see how necessitation implies strong supervenience, suppose that necessitation is true: for any object *x* in any possible world that has any *A*-property, *F*, there is a *B*-property, *G*, such that *x* has *G*. Moreover, if necessitation is true, then in any possible world, anything that has *G* has *F*. Suppose, then, that $F_1, F_2, \ldots, F_n$ are all the *A*-properties of *x*. According to necessitation, for each F_i that *x* has there will be a corresponding *B*-property, G_i, that *x* has as well. Suppose, then, that $G_1, G_2, \ldots, G_n$ are the *B*-properties corresponding to *x*'s *A*-properties: *x*'s having F_1 corresponds to *x*'s having G_1, *x*'s having F_2 corresponds to *x*'s having G_2, and so on. Given these correlations, necessitation implies that necessarily anything with G_1 will have F_1, anything with G_2 will have F_2, and so on. Consider now an object, *y*, in any possible world, that is *B*-indistinguishable from *x*. To say that *x* and *y* are *B*-indistinguishable is to say that for any *B*-property, G_i, *y* has G_i if and only if *x* does. Hence, *x* and *y* have all the same *B*-properties: both have $G_1, G_2, \ldots, G_n$. We saw already that according to necessitation, something has G_1 only if it has F_1, and it has G_2 only if it has F_2, and so on for every property G_i. Since *x* and *y* have all the same *B*-properties, it follows that they will have all the same *A*-properties as well. Moreover, this is true necessarily: it obtains in every possible world. Consequently, in whatever world *y* is in, *y* has *B*-property G_i only if it has the corresponding *A*-property, F_i. Since *x* and *y* have all the same *B*-properties, it follows that they will have all the same *A*-properties. Consequently, in whatever worlds they are in, *x* and *y* will be *A*-indistinguishable if they are *B*-indistinguishable. Hence, necessitation implies strong supervenience.

defended in Chapters 2–5 rejects it. That theory claims that properties are sparse (Section 2.1), and in that case, there is little reason to think that properties are closed under complementation. Brian McLaughlin puts the point this way:

> If one takes (empirical) properties to be ways things might be that endow them with causal powers, then one might deny… [that] for any *B*-property *P*, not having *P* is itself a *B*-property… For one might hold that not being a certain way is not a way of being that endows anything with causal powers. (1995: 28)

McLaughlin (1995: 29–30) goes on to show that even if one grants Kim's assumptions about properties, strong supervenience still does not imply necessitation.[3]

A hylomorphic theory committed to the embodiment thesis implies some robust necessitation theses for structured individuals. Here is one of them:

> *Hylomorphic activity necessitation*: Necessarily, for any structured individual, x, if x engages in activity A at time t, then (a) x has proper parts and surrounding materials with powers such that the A-wise manifestation of those powers at t composes x's A-ing at t, and (b) necessarily, for any individual z and time t^*, if z has proper parts and surrounding materials at t^* that are exactly similar to x's at t, then z engages in A-ing at t^*.

According to clause (a) of hylomorphic activity necessitation, any structured individual that engages in an activity, A, at a time has proper parts and surrounding materials with powers whose manifestations are structured A-wise. This is implied by a commitment to the embodiment thesis defended in Chapter 8. That thesis claims that all the powers of structured individuals are essentially embodied in their parts. The parts that embody a's power to A are the parts that in conjunction with surrounding materials would compose a's A-ing at t if those parts and surrounding materials were to manifest their powers A-wise at t. Consequently, if all of an individual's activities are essentially embodied, it follows that all of its activities will be composed of the structured manifestation of the powers of its parts and surrounding materials. As a result, if a engages in A-ing at time t, the powers of a's parts and surrounding materials will manifest themselves A-wise at t.

According to clause (b) of hylomorphic activity necessitation, if some x engages in A-ing at t, then necessarily, for any structured individual, z, whose proper parts and surrounding materials at time t^* are exactly similar to x's proper parts and surrounding materials at time t, z will engage in A-ing at t^*. The argument for clause (b) is this: (1) necessarily, if z's parts and surrounding materials at t^* are exactly similar to x's parts and surrounding materials at t, then those parts and surrounding materials

[3] Kim concedes that strong supervenience and necessitation may not be equivalent, but adds, "this is of little philosophical significance since there seem to be no interesting real-life applications in which one holds and the other doesn't" (2003: 561 n. 2).

have powers whose manifestations are structured A-wise at t^*. But (2) necessarily, if z's parts and surrounding materials have powers whose manifestations are structured A-wise at t^*, then the structured manifestation of those powers composes z's A-ing at t^*, and (3) necessarily, if the structured manifestation of the powers of z's parts and surrounding materials compose z's A-ing at t^*, then z engages in A-ing at t^*. Let us consider these premises one at a time.

Premise (1) follows from the notion of exact similarity. Necessarily, if a at time t and b at time t^* are exactly similar, then a has F at t if and only if b has F at t^*. a's F-ing at t and b's F-ing at t^* are exactly similar tropes. Suppose then that z's proper parts and surrounding materials at time t^* are exactly similar to x's proper parts and surrounding materials at time t. In that case, x's proper parts and surrounding materials will manifest their powers A-wise at t if and only if z's proper parts and surrounding materials manifest their powers A-wise at t^*.

Premise (2) follows from the definition of activity composition introduced in Section 8.2. According to that definition, if a is a structured individual with the power to engage in A-ing, and $b_1, b_2, \ldots, b_n$ are individuals, a subset of which are proper parts of a, each having a power to engage in an activity, A_i, then a's A-ing at t is composed of b_1's A_1-ing at t_1, b_2's A_2-ing at t_2, … and b_n's A_n-ing at t_n exactly if b_1's A_1-ing at t_1, b_2's A_2-ing at t_2, … and b_n's A_n-ing at t_n are structured A-wise, where time t includes the times $t_1, t_2, \ldots, t_n$. According to this definition, if a's proper parts in conjunction perhaps with other individuals (surrounding materials) manifest their powers A-wise during a time, then they compose a's A-ing at that time. The structured manifestation of the powers of something's parts and surrounding materials is sufficient to compose its activities.

Premise (3) claims that activity composition is sufficient for an individual to engage in an activity. The idea behind this claim is that there is nothing to engaging in an activity on the hylomorphic view other than imposing the right activity-making structure on the way something's parts and surrounding materials manifest their powers. In the same way, there is nothing to there being a structured individual on the hylomorphic view other than imposing the right individual-making structure on some materials. On the hylomorphic view, structure *matters*: it operates as a basic ontological principle. Structure also *counts*: it explains the unity of composite entities, including the unity of composite activities. If the A-wise manifestation of the powers of something's parts and surrounding materials at a time is not sufficient for that thing to engage in A-ing at that time, then nothing is sufficient for it on the hylomorphic view. On that view, there is nothing to something's A-ing other than its imposing A-wise structure on its parts and surrounding materials: for z to impose an A-wise structure on the way its parts and surrounding materials manifest their powers at t^* is precisely for z to engage in A-ing at t^*. Consequently, if the structured manifestation of the powers of z's parts and surrounding materials compose z's A-ing at t^*, then z engages in A-ing at t^*.

The embodiment thesis also implies a commitment to a necessitation thesis for the powers of structured individuals:

> *Hylomorphic power necessitation*: Necessarily, for any structured individual, x, if x has the power to engage in activity A at time t, then (a) x has proper parts with powers such that the A-wise manifestation of those powers in conjunction with the powers of surrounding materials at a time, t^*, would compose x's A-ing at t^*, and (b) necessarily, for any individual z and time t^*, if z has proper parts at t^* exactly similar to x's at t, then z has the power to engage in A-ing at t^*.

According to clause (a), any individual that has the power to engage in an activity, A, at a time has proper parts with powers such that the manifestations of those powers, in conjunction with those of surrounding materials, are capable of being structured A-wise. Like clause (a) of hylomorphic activity necessitation, this claim too is implied by a commitment to the embodiment thesis. The latter implies that if a has the power to engage in A-ing, then that power is essentially embodied in the proper parts of a; specifically, it is embodied in those parts that in conjunction with surrounding materials would compose a's A-ing at t if they and the surrounding materials were to manifest their powers A-wise at t.

Clause (b) of hylomorphic power necessitation says that if x has the power to engage in A-ing at t, then any structured individual, z, with proper parts exactly similar to x's at time t^* will also have the power to engage in A-ing; it implies that z's having such parts is sufficient for its having the power. The argument for clause (b) of hylomorphic power necessitation parallels the argument for clause (b) of hylomorphic activity necessitation: Necessarily, if z's proper parts at t^* are exactly similar to x's proper parts at t, then those parts have powers exactly similar to the powers of x's parts. If that is the case, then it is possible for the powers of z's parts, in conjunction with the powers of surrounding materials, to be structured A-wise, and if that is possible, then it is possible for the powers of those parts and surrounding materials to manifest themselves in a way that composes z's A-ing at a time. But on the hylomorphic view, to say that z has proper parts at time t^* whose powers in conjunction with those of surrounding materials can manifest themselves in a way that composes z's A-ing, is to say that z has the power to engage in A-ing at t^*. Recall that on the hylomorphic view there is nothing to engaging in an activity other than imposing the right activity-making structure on the way something's parts and surrounding materials manifest their powers. Consequently, if something has parts whose powers in conjunction with the powers of surrounding materials can be structured in the right way to compose its A-ing, then that thing has the power to engage in A-ing. Consequently, if x has proper parts that embody its power to engage in A-ing, then necessarily, anything that has exactly similar proper parts will have an exactly similar power.

9.2 Activity and Power Supervenience

The necessitation theses defended in Section 9.1 imply corresponding strong supervenience theses, provided that necessity is understood as quantification over possible worlds:

Hylomorphic activity supervenience: For any possible worlds w_1 and w_2, and any individual x in w_1 and individual y in w_2, if the activities of x at time t and y at time t^* are composed of the structured manifestations of the powers of exactly similar parts and surrounding materials, then y will engage in activities at t^* exactly similar to the activities of x at t.

Hylomorphic power supervenience: For any possible worlds w_1 and w_2, and any individual x in w_1 and individual y in w_2, if x at time t and y at time t^* have exactly similar parts and surrounding materials, then y will have powers at t^* exactly similar to the powers of x at t.

To see how hylomorphic activity necessitation implies hylomorphic activity supervenience, suppose that the former claim is true: For any structured individual x in any possible world, if x engages in any activity, A, then there are proper parts and surrounding materials with powers such that the A-wise manifestation of those powers at a time composes x's A-ing at that time. Moreover, if hylomorphic activity necessitation is true, then in any possible world, anything that has exactly similar parts and surrounding materials at a time, t^*, will engage in A-ing at t^*. Suppose, then, that $A_1, A_2, \ldots, A_n$ are all the activities in which x engages at time t. Hylomorphic necessitation implies, for each A_i that x engages in, that there will be corresponding parts and surrounding materials with powers that manifest themselves A_i-wise. In that case, x's engaging in A_1 corresponds to x's having parts and surrounding materials whose powers manifest themselves A_1-wise, and x's engaging in A_2 corresponds to x's having parts and surrounding materials whose powers manifest themselves A_2-wise, and so on. Given these correlations, necessitation implies that necessarily anything having parts and surrounding materials whose powers manifest themselves A_1-wise will engage in A_1-ing, and anything having parts and surrounding materials whose powers manifest themselves A_2-wise will engage in A_2-ing, and so on. Consider now an object, y, in any possible world, whose parts and surrounding materials are exactly similar to x's. To say that x and y have exactly similar parts and surrounding materials is to say that for any A_i-wise manifestation of the powers of those parts and surrounding materials, y's parts and surrounding materials manifest themselves A_i-wise if and only if x's parts and surrounding materials do. Hence, x's and y's parts and surrounding materials will all manifest themselves in exactly similar ways: both have parts and surrounding materials that manifest their powers A_1-wise, A_2-wise, $\ldots$, A_n-wise. We saw already that according to hylomorphic necessitation, something has parts and surrounding

materials that manifest their powers A_1-wise only if it engages in A_1-ing, and it has parts and surrounding materials that manifest their powers A_2-wise only if it engages in A_2-ing, and so on for every A_i-wise structure. Since all of x's and y's parts and surrounding materials are exactly similar, it follows that x and y will engage in exactly similar activities. Moreover, this is true necessarily: it obtains in every possible world. Consequently, for whatever worlds, w_1 and w_2, that x and y are in, x in w_1 and y in w_2 will engage in exactly similar activities if they have exactly similar parts and surrounding materials. Hylomorphic activity necessitation thus implies hylomorphic activity supervenience.

A similar argument shows that hylomorphic power necessitation implies hylomorphic power supervenience. Suppose that the former claim is true: For any structured individual x in any possible world that has the power to engage in any activity, A, there are proper parts with powers such that if those powers and the powers of surrounding materials manifest themselves A-wise at a time, they will compose x's A-ing at that time. Moreover, if hylomorphic power necessitation is true, then in any possible world, anything that has exactly similar parts at a time will have the power to engage in A-ing at that time. Suppose, then, that $A_1, A_2, \ldots, A_n$ are all the activities in which x has the power to engage at time t. Hylomorphic power necessitation implies that for each A_i in which x has the power to engage there will be corresponding parts with powers that together with surrounding materials can manifest themselves A_i-wise. In that case, x's having the power to engage in A_1 corresponds to x's having parts whose powers in conjunction with the right surrounding materials can manifest themselves A_1-wise, and x's having the power to engage in A_2 corresponds to x's having parts whose powers in conjunction with the right surrounding materials can manifest themselves A_2-wise, and so on. Given these correlations, necessitation implies that necessarily anything with parts whose powers in conjunction with the powers of surrounding materials that can manifest themselves A_1-wise will have the power to engage in A_1-ing, and anything with parts whose powers in conjunction with the powers of surrounding materials that can manifest themselves A_2-wise will have the power to engage in A_2-ing, and so on. Consider now an object, y, in any possible world, whose parts are exactly similar to x's. To say that x and y have exactly similar parts and surrounding materials is to say that for any A_i-wise manifestation of which those parts in conjunction with surrounding materials are capable, y's parts will have that capability if and only if x's parts do. Hence, x's and y's parts will all be capable of manifesting themselves in conjunction with surrounding materials in exactly similar ways: both have parts that are capable of manifesting their powers in conjunction with surrounding materials A_1-wise, A_2-wise, $\ldots$, A_n-wise. We saw already that according to hylomorphic power necessitation, something has parts that are capable of manifesting their powers in conjunction with surrounding materials A_1-wise only if it has the power to engage in A_1-ing, and it has parts that are capable of manifesting their powers in conjunction with surrounding materials A_2-wise only if it has the power to engage in A_2-ing, and

so on for every A_i-wise structure. Since all of *x*'s and *y*'s parts and surrounding materials are exactly similar, it follows that *x* and *y* will have the power to engage in exactly similar activities. Moreover, this is true necessarily: it obtains in every possible world. Consequently, for whatever worlds, w_1 and w_2, that *x* and *y* are in, *x* in w_1 and *y* in w_2 will have exactly similar powers if they have exactly similar parts. Hylomorphic power necessitation thus implies hylomorphic power supervenience.

Several remarks are in order about hylomorphic necessitation and supervenience. First, it will be helpful to categorize them in terms of the standard distinctions that have been used to describe supervenience and necessitation relations (Kim 1984; 1987; 1988; 1990; McLaughlin 1995). First, hylomorphic necessitation and supervenience are relations between individuals in multiple domains, not individuals in a single domain. The supervenient domain consists of composite wholes with their powers and activities, while the subvenient domain consists of the parts and surrounding materials of those wholes with their powers and activities. In addition, hylomorphic necessitation and supervenience are based on indiscernibility or exact similarity, not on a looser notion of similarity, and they are synchronic relations, not diachronic ones.

Necessitation and supervenience relations are also categorized in terms of their subvenient or base conditions. Supervenience bases are often described as local or global. Roughly, if *A*-properties supervene on *B*-properties, the base for that supervenience relation is local if the *A*-properties that an individual has covary simply with the *B*-properties of that individual or perhaps its parts. The base is global, on the other hand, if the *A*-properties of individuals across an entire world covary with the *B*-properties of individuals across that world. Hylomorphic necessitation and supervenience bases, however, are not easily categorized in these terms. Suppose that *a* is an individual that engages in activity *A* at time *t*, and that the *y*s are individuals with powers $P_1, P_2, \ldots, P_m$ whose structured manifestation composes *a*'s *A*-ing at *t*. In that case, the necessitation base and the supervenience base for *a*'s *A*-ing at *t* is the *y*s' manifesting $P_1, P_2, \ldots, P_m$ *A*-wise at *t*. The necessitation and supervenience bases for an activity might include only individuals that are parts of *a*, or they might include individuals that are parts of *a* and that are in *a*'s immediate environment, or they might include individuals that are much farther away than what we would typically consider *a*'s immediate environment. If Gabriel sees Polaris in the night sky, then arguably Polaris (the distant physical materials that are spatially arranged star-wise) are included among the individuals whose activities compose Gabriel's seeing, and hence those materials would be included in the necessitation/supervenience base for that activity. Tyler Burge makes a similar point about thinking:

> Some of the arguments for anti-individualism indicate that the supervenience bases of empirical thoughts are not local to the body of the individual. The supervenience base of a thought about mercury and Lake Baakal would involve a complex pattern of individual-environment

relations, including causal-perceptual relations to objects in the environment. The supervenience base of such thoughts is a massively complex pattern that includes states in the thinker's body at the time of the thought, but extends over large stretches of space and time. Such a pattern is not local to the individual's body. (2006: 374)

When it comes to hylomorphic necessitation and supervenience bases, it is also important to remember that the individuals whose activities compose Gabriel's seeing include individuals all the way down the compositional hierarchy. Gabriel can see only if his eyes are working properly, but his eyes can work properly only if the cells composing their retinas are working properly, and those cells can work properly only if various other lower-level conditions are satisfied. All of the individuals, powers, and manifestations in this compositional hierarchy, including those at a fundamental physical level, belong to the necessitation/supervenience base for Gabriel's seeing.

In general, the necessitation and supervenience bases for the activities of any given individual include all those individuals, powers, and manifestations that are necessary to compose its activities.

The full list of social, environmental, physiological, and other conditions that are necessary for something's activities are typically not discoverable a priori. They are instead discovered a posteriori through methods like functional analysis. Those conditions, moreover, are necessary not just nomologically, but metaphysically. The reasons for this were touched on earlier (Sections 8.1–8.2): the activities of structured individuals are essentially composite; they involve those individuals, powers, and manifestations which are included in their necessitation and supervenience bases. The modal operators in the definition of hylomorphic necessitation should thus be taken to express metaphysical necessity. Likewise, the quantifiers in the definition of hylomorphic supervenience should be taken to range over metaphysically possible worlds.

9.3 Structo-Physical Necessitation and Supervenience

I have argued so far in this chapter that the embodiment thesis implies that the activities and powers of structured individuals are necessitated by and supervenient upon the powers of their parts and the structured manifestation of those powers. Consider now necessitation and supervenience theses of a different sort—theses according to which structured individuals, their powers, and activities are necessitated by and supervene upon spatial arrangements of physical materials. To bring these theses into focus, imagine an individual such as Godehard, who is composed of physical materials, $p_1, p_2, \ldots, p_n$, that are structured human-wise. Imagine now that there are physical materials, $q_1, q_2, \ldots, q_n$, exactly similar to the *p*s. The *q*s are arranged spatially in just the way the *p*s are, and are exactly similar in respect of all the properties and relations that can be described exhaustively by physics. Is the exact similarity of the *q*s to the *p*s sufficient to ensure that the *q*s, like the *p*s, compose an

individual exactly similar to Godehard—a human being (or human clone) that has powers to engage in activities exactly similar to his?[4] If the answer is yes, then this suggests that claims like the following are true:

> *Structo-physical necessitation*: Necessarily, for any x, if x is a structured individual at time t, then (a) there are physical materials, $y_1, y_2, \ldots, y_n$, that compose x at t, and (b) necessarily, for any physical materials, $z_1, z_2, \ldots, z_n$, and time t^*, if the zs at t^* are exactly similar to the ys at t in respect of those properties and relations that can be exhaustively described by physics, then they compose an individual at t^* that is structurally exactly similar to x at t.
>
> *Structo-physical supervenience*: For any possible worlds w_1 and w_2, and any physical materials, $x_1, x_2, \ldots, x_n$, in w_1 and $y_1, y_2, \ldots, y_n$ in w_2, if the xs at time t are exactly similar to the ys at time t^* in respect of the kinds of properties and relations that can be exhaustively described by physics, then the xs compose an individual at t if and only if the ys compose an individual at t^* that is structurally exactly similar to x at t.

The hylomorphic necessitation and supervenience theses discussed in Sections 9.1 and 9.2 were implications of the embodiment thesis defended in Section 8.3. That thesis implies that the powers of structured individuals are essentially embodied in physical parts of some sort, but it does not say whether the existence of structured individuals themselves is something that is necessitated by or supervenient upon spatial arrangements of physical materials. These claims require independent argument.

Consider first structo-physical necessitation. According to clause (a), if there is a structured individual at time t, then there must be physical materials that compose that individual at t. This seems to be implied by the claims that *structure matters* and that *structure counts*. According to the former principle, structure is a basic ontological principle; according to the latter, structure is responsible for unifying diverse physical materials into a single unified whole. Structured individuals are essentially organizers or configurers of materials (Section 6.1), and composition occurs when and only when an individual configures materials (Sections 6.3–6.4). There is an x such that the ys compose x if and only if x is an individual that configures the ys. The right-to-left implication is what's relevant for our purposes here: if there is a structured individual, x, that configures the ys, then the ys compose x. This is clause (a) of structo-physical necessitation.

Now, given that x is a structured individual which is composed of the ys, clause (b) says that necessarily, if there are zs exactly similar to the ys in respect of those characteristics which can be described exhaustively by physics, then the zs compose an individual that is structurally exactly similar to x. The argument for clause (b) is

4 I add the disjunct 'or human clone' in case being a genuine human requires a historical condition such as having come into existence by a process of natural selection.

this: (1) necessarily, if the *z*s at t^* are exactly similar to the *y*s at t in respect of the kinds of properties and relations that can be described exhaustively by physics, then the *z*s must be structured at t^* in a way that is exactly similar to the way the *y*s are structured at t. But (2) necessarily, if the *z*s are structured at t^* in a way exactly similar to the way the *y*s are structured at t, then the *z*s compose an individual that is structured exactly similar to x at t^*. I'll treat these premises in reverse order.

Premise (2) follows from the hylomorphic account of composition discussed in Chapter 6. According to that account, composition happens if physical materials come to have an individual-making structure. To deny that having an individual-making structure is sufficient for physical materials to compose a structured individual is either to deny that *structure matters* or to deny that *structure counts*; it is either to deny that hylomorphic structure operates as a basic ontological principle, one that, in conjunction with physical materials, accounts for what things are, or else it is to deny that structure is what unifies those materials so that they compose a whole. If, however, *structure matters* and *structure counts*, then physical materials will compose an individual, provided only that they are structured in some way.

Premise (1) is the real crux of the argument. It claims that if the *y*s are structured in a certain way and the *z*s are physically exactly similar to the *y*s, then those physical similarities by themselves are sufficient to ensure that the *z*s will be structured in an exactly similar way. One argument in favor of this claim considers the implications if it is false. If Premise (1) is false, then it is possible that the *z*s might be physically exactly similar to the *y*s, and yet not be structured the same way. Suppose, for instance, that Godehard is composed of physical materials, $p_1, p_2, \ldots, p_n$. If Premise (1) is false, then it is possible for there to be physical materials, $q_1, q_2, \ldots, q_n$, which are exactly similar to the *p*s in respect of the properties and relations that can be described exhaustively by physics, but which, unlike the *p*s, are not structured human-wise. Naturally, because the *q*s have physical characteristics exactly similar to the *p*s—including exactly similar spatial arrangements—it might appear to anyone who observes them that the *q*s compose an exact replica of Godehard—a human being (or human clone) that has powers and engages in activities exactly similar to his. But if Premise (1) is false, such an appearance can be misleading. In the case we are considering, because the *q*s are not structured human-wise, they do not compose a human being, nor do they embody any distinctively human powers or compose any distinctively human activities. If Premise (1) is false, this could happen in either of two ways: either the *q*s have a structure, just not a human one, or else the *q*s are not structured in any way at all. In the latter case, the *q*s would be an instance of what we might call a *structure zombie*: they would be physical materials that in every respect appeared to compose a structured individual but which did not compose any structured individual at all. I regard both of these cases as absurd, but I confess that I do not have an argument to show that they are impossible. I can only hope to appeal to the relevant intuitions.

It will help to explain why I find the case of structure zombies absurd by contrasting it with a similar case that I do not find absurd. Imagine a very powerful being—a demiurge—who is able to gather together the *q*s and manipulate them in such a way that they are arranged spatially in just the way the *p*s are. The demiurge, however, uses its power to ensure that the *q*s do not interact with each other in the way that the *p*s do. The *q*s, for instance, do not form any of the relevant atomic or other bonds that the *p*s do; rather, these bonds and other physical relations are merely simulated by the demiurge, who micromanages the various changes that the *q*s undergo. As in the structure zombie case, anyone who observes the demiurge-manipulated *q*s might think that they compose an exact replica of Godehard, and as in the structure zombie case, such an appearance is misleading. What distinguishes this case from the structure zombie case is that the *q*s are not exactly similar to the *p*s in respect of the kinds of properties and relations that can be described exhaustively by physics, for in this case, the demiurge interferes with the ways in which the *q*s interact.

Unlike the structure zombie case, I do not find the case of demiurgical manipulation absurd. What the demiurge does is interfere with the natural course of physical events. It prevents the *q*s from interacting with each other in the ways that the *p*s do. It effectively uses the *q*s as a puppeteer would: it manipulates them to mimic changes that the *q*s would undergo were it not for the demiurge's interference. I do not find it absurd that a being might possess and manifest powers of this sort. Nor do I find it absurd that a demiurgical being might jump-start, so to speak, the existence of a structured individual by manipulating the *q*s in such a way that through their natural interactions they should come to be structured human-wise. In fact, I think it might be possible in the future for humans themselves to be capable of manipulating physical materials in just such a way. What I find absurd is instead the idea that the *q*s might manifest all their own powers and interact with each other physically in exactly the ways that the *p*s do, and yet not come to be structured human-wise. I think that physical materials have powers to be structured in various ways, and that they manifest these powers on their own under the right conditions. If this is the case, then provided that the *q*s are exactly similar to the *p*s, and that they interact with each other physically in exactly the ways the *p*s do, then the *q*s will be structured human-wise exactly if the *p*s are. This would rule out structure zombies.

The other case that would challenge Premise (1) is a case in which the *q*s are physically exactly similar to the *p*s, and the *q*s interact with each other physically in exactly the way the *p*s do, but in which the *q*s are structured in some nonhuman way. I regard this case as absurd as well. The following line of reasoning will perhaps explain why. Godehard is composed of the *p*s, which are structured human-wise. Because the *q*s also have a hylomorphic structure, they compose an individual as well—call him 'Godehard*.' It seems plausible to think that different kinds of structures must confer different kinds of powers on the individuals having them. This seems to be an implication of the identity conditions for powers sketched in

Section 4.4. If that is the case, and Godehard*'s structure is different from Godehard's, then that structure must confer on Godehard* powers that differ in kind from Godehard's. According to the embodiment thesis, all of an individual's powers are essentially embodied. Consequently, because Godehard and Godehard* have different kinds of powers, they must also have different kinds of parts at some level. The account of composition defended in Chapter 6 implies that parts differ from each other if and only if they contribute to the wholes they compose in different ways. So for the *p*s and the *q*s to differ from each other in kind, they would have to contribute to the activities of the higher-level things they compose in different kinds of ways. But it is difficult to see how the contributions of Godehard's and Godehard*'s parts could differ at any level.

To appreciate the difficulty of imagining Godehard's and Godehard*'s parts contributing to their respective activities in different ways, let us suppose that the *p*s and the *q*s belong to the fundamental physical level of composition. For simplicity, let us call the next higher level of composition the 'atomic level,' and let us call the higher-level wholes composed of the *p*s and the *q*s at that level 'atoms.' In order for the *p*s and the *q*s to qualify as different kinds of parts, they would have to contribute to the activities of the atoms they compose in different kinds of ways. But it is difficult to see how this could be the case if the *p*s and the *q*s are exactly similar in all the respects that can be described by fundamental physics. The only way their contributions could differ, it seems, is if the activities of Godehard's atoms and Godehard*'s atoms themselves engaged in different kinds of activities. Someone might suppose that exactly similar fundamental physical interactions among the *p*s and the *q*s could contribute to one kind of activity in Godehard's atoms, and another kind of activity in Godehard*'s. In fact, someone might argue that by analogy exactly similar arm movements could contribute to greeting someone in one circumstance, and to hailing a cab in another. But it is difficult to see how the envisaged analogy could obtain in the case of the *p*s and the *q*s. The difference between greeting someone and hailing a cab depends on differences in the surrounding environments in which the two activities occur—differences which themselves depend on lower-level physical ones. The proximity of a cab in the one case, and of a person who can be greeted in the other, depend on differences at a fundamental physical level. But in the case of the atoms, there are no differences at a fundamental physical level: the *p*s and the *q*s are, by hypothesis, exactly similar in all the respects that can be described by fundamental physics. How, then, could either the powers or the activities of Godehard's and Godehard*'s atoms differ from each other if they are composed of fundamental physical parts that are exactly similar in all those respects? Again, if there were a difference in the powers of the atoms, then that difference would have to track some difference in their fundamental physical parts since the powers of things are essentially embodied. But there do not seem to be any differences in their fundamental physical parts since the *p*s and the *q*s are by hypothesis exactly similar.

What is true of the atomic level seems true of the next level in the composition hierarchy as well—call it the 'molecular level,' and call the wholes it comprises 'molecules.' Can Godehard's molecules differ from Godehard*'s? If Godehard's atoms are exactly similar to Godehard*'s, it is difficult to see how they could, for the powers of the molecules will be essentially embodied in their parts and the parts of the molecules are atoms. Because Godehard's and Godehard*'s atoms are exactly similar, it is difficult to see how they could embody different powers. The same reasoning can be applied mutatis mutandis, it seems, at every level in the compositional hierarchy. The powers of a whole at level$_N$ in the hierarchy are embodied in parts at level$_{N-1}$, but for each level$_N$, it seems that Godehard's parts at level$_{N-1}$ are exactly similar to Godehard*'s parts at level$_{N-1}$, and consequently, it is difficult to see how those parts could embody different powers. Given the hylomorphic view I've defended, then, it is difficult to see how Godehard's and Godehard*'s parts could contribute to Godehard's and Godehard*'s respective activities in different kinds of ways, and because of that, it is difficult to see how Godehard and Godehard* could have different kinds of parts. But if they do not have different kinds of parts, it is difficult to see how they could have different kinds of powers, and if they do not have different kinds of powers, it is difficult to see how they could have different kinds of structures.

One objection to this line of reasoning is inspired by Mark Johnston (2006), who has argued that it is possible for the same physical materials to have different principles of unity. His argument appeals to several examples of objects which appear to be different but which are nevertheless composed of exactly the same parts. His first example is a physical system composed of superdense spheres, which are held together in a cube-like shape by magnetic forces and also by gravitational ones. According to Johnston, the very same spheres compose two different physical systems, one whose principle of unity is magnetic, the other whose principle of unity is gravitational. Johnston's second example is a spork, a camping implement that has fork-like tines at one end and a spoon-like bowl at the other. The materials composing the spork have three different principles of unity according to Johnston, a spoon-like principle, a fork-like principle, and a spork-like principle. The same materials thus compose three different objects: a spoon, a fork, and a spork. Johnston's third example involves a kind of tumbleweed that flowers only after a flooding rain. Some tumbleweeds of this kind die without ever flowering, and so consist of only a spherical stem structure. Yet, the tumbleweed is not identical to the stem structure, for if the tumbleweed were to blossom after a flooding rain, the flower would be a part of it but not part of the stem structure. The unflowered tumbleweed thus has two different principles of unity, the tumbleweed principle and the stem principle, which correspond to two different individuals: the tumbleweed and its stem. Johnston's final example involves an ambiguous road sign which reads 'Begin Highway':

Between Haifa and Tel Aviv a highway is to be built . . . The friends of Menachem Begin wish to memorialize Begin by naming the highway after him. So they pay a Mr. Janus . . . to put up a

> green and white sign naming the highway the Begin Highway . . . Janus is the very same man contracted by the highway builders to put up the usual signage around the highway. And this calls for . . . a green and white sign indicating the start of the highway, a sign saying . . . "Begin Highway" . . . So Janus forms the plan to use the very same piece of signage . . . to both name the highway after Begin and to indicate where the highway begins. He thus produces two signs in the same place at the same time . . . The one sign . . . names the highway . . . The other sign . . . indicates where the highway begins . . . The friends of Begin drive by and only see the first sign. Most other motorists only see the second. (2006: 669–70)

Johnston's examples purport to show that it is possible for numerically the same physical materials to have different principles of unity, and thus for those materials to compose numerically different individuals. I will take it as a given that if this is possible for numerically the same physical materials, then it is also possible for exactly similar physical materials, such as the *p*s and *q*s in the example of Godehard and Godehard*. Johnston's examples thus pose a challenge to the line of reasoning I've sketched.

The first thing to say in response is that principles of unity on Johnston's view are not the same as structures on the hylomorphic view I've defended. One difference is that *structures make a difference* on the hylomorphic view I've defended: they confer distinctive causal powers on their possessors. It's not clear that the same is true of Johnston's principles of unity. It is not clear, for instance, that the spoon, the fork, and the spork have different powers despite having different Johnstonian principles of unity. Closely related to this difference is another: structures are sparse, but Johnston's principles of unity are abundant, or at least more abundant than structures. As a result, his view leads to a proliferation of material objects, a topic to which I'll return in Chapter 14.

With these differences in mind, consider Johnston's examples. According to the hylomorphic view I've defended, there are no spoons, forks, or sporks. There are only physical materials that are spatially arranged spoon-wise, fork-wise, or spork-wise. Which spatial arrangement we choose to focus on in any given situation is a function of our descriptive, practical, and other aims in that situation. In focusing on the fork end or in using it, we impose a certain kind of activity-making structure on the way our parts and the spork-wise materials manifest their powers. That arrangement is different from the activity-making structure we impose when we focus on or use the spoon end, or when we take in the sum of materials that are spatially arranged spork-wise. These differences do not correspond to different spoon, fork, and spork structures; they correspond to different ways in which individuals like us impose activity-making structures on our parts and surrounding materials.

Something analogous is true of the road sign and the spheres. Whether people see the sign one way or the other is a function of their interpretive activities. Begin's friends and the motorists impose different activity-making structures on their parts and the physical materials that are spatially arranged road-sign-wise. Similarly, the superdense spheres do not compose two individuals, they compose nothing at all.

They are simply spatially arranged cube-wise, and the persistence of that spatial arrangement is overdetermined by the magnetic and gravitational powers the spheres manifest (or more precisely, by the magnetic and gravitational powers of the physical materials that are spatially arranged sphere-wise).

The case of the tumbleweed is different. The hylomorphic view I defend treats it in a manner similar to the brainstem case discussed in Section 7.1. It denies that the tumbleweed has a stem before it blossoms. Suppose that the tumbleweed is composed of objects $a_1, a_2, \dots, a_n$—the cells, organelles, molecules, and so on that would be revealed through functional analysis of the tumbleweed's activities. If the tumbleweed were to blossom at time *t*, then at that time the *a*s would come to compose a stem, but unless the tumbleweed blossoms, the *a*s compose only the tumbleweed; they do not compose a stem. Critics might complain that it is implausible to suppose that the tumbleweed's stem would come into existence exactly when some other part (the flower) does. Recall, however, that on the hylomorphic view, parthood depends on the contributions something makes to the activities of the whole to which it belongs. We saw in the brainstem case that those contributions could vary depending on which other parts exist. If a brainstem contributes to a whole organism by coordinating the activities of other parts, then plausibly there is no brainstem if there are no other parts. The tumbleweed's stem might depend on other parts as well. Alternatively, a hylomorphist could claim that there is no stem at all, that even after the tumbleweed flowers its only parts are the *a*s and its flower. What position a hylomorphist ought to take here should depend on broadly empirical considerations. The important point vis-à-vis Johnston's argument is that according to the hylomorphic view I've defended, the *a*s do not compose two different individuals with two different kinds of structures; there is only one: the tumbleweed.

I hope these remarks go some of the way toward explaining why I find it absurd that the *p*s and the *q*s might be physically exactly similar to each other and yet compose wholes that are structured in different kinds of ways. Again, I do not purport to have shown that it is impossible for them do so. Perhaps there is a hylomorphic line of reasoning that could show this, but I do not think I have provided one here. I have only tried to explain the intuitions supporting Premise (1). If those intuitions carry the day, then there are neither structure zombies nor cases in which exactly similar physical materials might be structured in different kinds of ways. In that case, however, there are no counterexamples to Premise (1), no cases in which the *q*s are physically exactly similar to the *p*s but in which the *q*s are not structured human-wise. And from Premise (1) and Premise (2), clause (b) of structo-physical necessitation follows.

I have explained why I find structo-physical necessitation to be plausible. But structo-physical necessitation implies structo-physical supervenience. The latter says that if the *x*s at *t* are exactly similar to the *y*s at t^* in respect of the kinds of properties and relations that can be exhaustively described by physics, then the *x*s compose an individual at *t* if and only if the *y*s compose an individual at t^* that is structurally

exactly similar to x. Suppose, then, that Godehard inhabits possible world_1 and Godehard* inhabits possible world_2, and that the *p*s composing Godehard at time t are exactly similar to the *q*s composing Godehard* at time t^* in respect of the kinds of properties and relations that can be exhaustively described by physics. Because the *p*s compose an individual that is structured human-wise at t, it follows from structo-physical necessitation that the *q*s must compose an individual that is structured human-wise at t^* as well. Moreover, the same reasoning works in the converse direction: because the *q*s compose an individual that is structured human-wise at t^*, it follows from structo-physical necessitation that the *p*s must compose an individual that is structured human-wise at t. The *p*s thus compose an individual that is structured human-wise at t if and only if the *q*s compose an individual that is structured human-wise at t^*. Structo-physical necessitation thus implies structo-physical supervenience.

9.4 Conclusion

A hylomorphic theory which is committed to the embodiment thesis defended in Chapter 8 implies some robust supervenience and necessitation theses. According to hylomorphic activity necessitation, if Gabriel engages in activity A at time t, there must be parts of him and perhaps surrounding materials with powers whose structured manifestations compose his A-ing at t. Necessarily, moreover, any individual, y, that has parts and surrounding materials at time t^* exactly similar to Gabriel's, will engage in A-ing at t^*. In addition, according to hylomorphic power necessitation, if Gabriel has the power to engage in A-ing at time t, there must be parts of him with powers such that the A-wise manifestation of those powers and the powers of surrounding materials at a time t^* would compose his A-ing at t^*, and necessarily, any individual, y, that has parts and surrounding materials exactly similar to Gabriel's at a time, has the power to engage in A-ing at that time. These necessitation theses imply corresponding strong supervenience theses. First, for any possible worlds w_1 and w_2, and any x in w_1 and y in w_2, if the activities of x at t and y at t^* are composed of the structured manifestations of the powers of exactly similar parts and surrounding materials, then y will engage in activities at t^* exactly similar to the activities of x at t. Second, for any possible worlds w_1 and w_2, and any x in w_1 and y in w_2, if x at t and y at t^* have exactly similar parts and surrounding materials, then y will have powers at t^* exactly similar to the powers of x at t.

I've also argued that structo-physical necessitation and supervenience are plausible theses for hylomorphists to endorse. Roughly, structo-physical necessitation says that necessarily, if x is a structured individual, then there are physical materials that compose x, and necessarily, for any materials that are physically exactly similar to the materials composing x (exactly similar, that is, in respect of those properties and relations that can be described exhaustively by physics), those physical materials compose an individual that is structurally exactly similar to x. Structo-physical

necessitation implies structo-physical supervenience: roughly, for any possible worlds w_1 and w_2, and any physical materials, $x_1, x_2, \ldots, x_n$, in w_1 and $y_1, y_2, \ldots, y_n$ in w_2, if the *x*s are physically exactly similar to the *y*s, then the *x*s compose a structured individual if and only if the *y*s compose an individual that is structurally exactly similar. Why endorse structo-physical necessitation? Denying it has the highly counterintuitive implication that it is possible that the *y*s might be physically exactly similar to the *x*s, and yet not be structured in an exactly similar way. There are two possible cases of this sort: either the *y*s have a structure, just not an exactly similar one, or else the *y*s are not structured in any way at all. In the latter case, the *y*s would be an instance of a *structure zombie*, physical materials that appear in every respect to compose a structured individual but that do not compose any structured individual at all. In the former case, the *y*s would be physically exactly similar to the *x*s, but would compose an individual with a different kind of structure. I consider both cases to be highly counterintuitive for reasons I described in detail in Section 9.3, but I do not purport to have shown that either case is impossible.

10

Explanation and Lower-Level Determination

10.1 Necessitation, Supervenience, and Explanation

One of the most common misunderstandings about the nature of necessitation and supervenience relations is that they entail some type of explanatory condition. On this misguided view, saying that A-properties are necessitated by or supervene upon B-properties implies that B-properties explain A-properties.

To appreciate why this view of supervenience and/or necessitation is misguided, consider again the general definition of necessitation introduced in Section 9.1. It says that necessarily for any x, if x has an A-property, F, then there is a B-property, G, such that x has G, and necessarily, for any z, if z has G, then z has F. What this definition does not say is that z's having G *explains* its having F, or that x has F *because* it has G. Someone might argue that metaphysically necessary property correlations always imply explanatory conditions, but this is not part of the definition of necessitation either. It represents a further condition that must be added.

What is true of necessitation is also true of strong supervenience. Because necessitation implies strong supervenience, it must imply anything that strong supervenience does. If strong supervenience included a condition that necessitation did not, that condition would not be derivable from necessitation by itself and the implication would not hold. Consequently, if necessitation doesn't imply an explanatory condition, strong supervenience cannot imply an explanatory condition either. On its face, then, the claim that necessitation and/or supervenience implies an explanatory condition is inaccurate; it freights in something that simply isn't there.

The point is especially clear if we contrast four views. According to each, mental properties are necessitated by and supervene upon physical properties, but what explains the necessitation and supervenience relation in each case is different.

A. The psychophysical identity theory claims that every mental property is identical to a physical property. This implies both psychophysical necessitation and psychophysical supervenience. Necessarily for every mental property, M, there is a physical property, P, such that anything that has M also has P. Moreover, since P is identical to M, necessarily anything that has P also has M. Mental properties are thus necessitated by physical properties. In addition, for every

mental difference between an individual x and an individual y there must be a corresponding physical difference, for if mental properties are identical to physical properties, every mental difference just is a physical difference. Mental properties thus supervene on physical properties.

B. A theory that grounds psychophysical parallelism in divine pre-established harmony can imply psychophysical necessitation and supervenience. Suppose that God sets up all possible worlds in such a way that for every mental property, M, there is a corresponding physical property, P, such that if x has M, then x has P, and necessarily any y that has P also has M. In that case, mental properties are necessitated by physical properties. Moreover, since necessitation implies supervenience, it follows that necessarily there will be a mental difference between individuals x and y only if there is a physical difference between them.

C. An emergentist or epiphenomenalist theory that is committed to a dualism of mental and physical properties can imply psychophysical necessitation and supervenience. Imagine that there are brute psychophysical laws, according to which necessarily for every mental property, M, there is a physical property, P, such that anything that has M must also have P, and necessarily everything with P must have M. Again, since necessitation implies supervenience, it follows that necessarily there will be mental differences between individuals only if there are physical differences between them.

D. Psychophysical eliminativism claims that there are no mental properties. Imagine a radical eliminativist theory, according to which the very concept of a mental property is incoherent. Not only do mental properties not exist on this view, it is impossible for them to exist. On such a view, psychophysical necessitation and supervenience follow as trivial implications. Since there can be no mental properties, the antecedent of necessitation ends up being false, so psychophysical necessitation ends up being true: since necessarily no x has mental properties, it follows that necessarily, for any x, if x has a mental property, M, then there is a physical property, P, such that x has P, and necessarily anything that has P must also have M. Likewise, since there can be no mental properties, on this view, it follows that there can be no mental differences. It thus trivially follows that necessarily if x and y differ mentally, they differ physically.

Each of these views is committed to psychophysical necessitation and supervenience, yet what explains the necessitation and supervenience relations on each view is different. On the first view, what explains the relations is a commitment to psychophysical identities; on the second, it is God's pre-established harmony; on the third, there is no explanation, each psychophysical correlation is taken as a brute law; and on the fourth view, psychophysical necessitation and supervenience are trivial implications of the claim that the concept of a mental property is incoherent.

Necessitation and supervenience are both compatible with a wide range of explanatory relations. Neither implies a particular account of why one kind of property necessitates or supervenes on another. In fact, neither implies that there be any such account at all; both are compatible with taking necessitation and supervenience as brute matters of fact.

The key insight here is that necessitation and supervenience merely state patterns of property covariation: they do not purport to explain why the patterns obtain, nor do they imply that higher-level conditions obtain because of lower-level ones. The point is not new. It has been in fact a central premise in criticisms of nonreductive physicalism. Like all physicalist theories, nonreductive physicalism is committed to the claim that everything can in principle be described and explained exhaustively by physics (a point I discuss in greater detail in Chapter 11). It nevertheless denies that psychological categories are correlated in a systematic way with physical ones. Since psychophysical reduction requires systematic psychophysical correlations, nonreductive physicalism implies that psychological discourse is not reducible to physics. A physicalist theory must nevertheless place some constraints on how mental properties covary relative to physical properties. The reason is that physicalism implies that everything depends in some sense on the physical. If the mental were able to vary relative to the physical ad libitum, it is not clear in what sense the mental could be said to depend on the physical. The burden for nonreductive physicalists is thus to posit a type of psychophysical dependence relation that falls short of providing the kinds of systematic psychophysical correlations that sanction psychophysical reduction. During the 1980s, many nonreductive physicalists felt confident that necessitation or supervenience could provide the needed dependence relation. The problem, as Jaegwon Kim (1984, 1990), Thomas R. Grimes (1988), and others argued, is that neither necessitation nor supervenience implies dependence. For one thing, relations like dependence and explanation are asymmetrical: if A-properties depend on or explain B-properties, it is not the case that B-properties depend on or explain A-properties. But necessitation and supervenience imply no such asymmetry. Kim described the situation:

> Much of the philosophical interest that supervenience has elicited lies in the hope that it is a relation of dependency; many philosophers saw in it the promise of a new type of dependency relation that seemed just right, neither too strong nor too weak, allowing us to navigate between reductionism and outright dualism. And it is the dependency aspect of supervenience, not the covariation component, that can sanction many of the usual philosophical implications drawn from, or associated with, supervenience theses concerning various subject matters. Often it is thought, and claimed, that a thing has a supervenient property *because*, or *in virtue of the fact that*, it has the corresponding base property, or that its having the relevant base property *explains* why it has the supervenient property. All these relations are essentially asymmetric . . . Clearly, property covariation by itself does not warrant the use of "because", "in virtue of", etc., in describing the relationship any more than it warrants the attribution of dependence . . . Property covariation *per se* is metaphysically neutral. (Kim 1990: 147–8)

Despite work like Kim's, misunderstandings about necessitation, supervenience, and their relations to explanation have persisted. It is worth noting, then, that the hylomorphic theory I've described is committed to the powers and activities of structured individuals being necessitated by and supervening upon the powers and activities of their parts and surrounding materials, but it is not committed to the claim that the powers and activities of those parts and surrounding materials explain the powers and activities of structured wholes. In the remainder of this chapter I explain why.

10.2 Lower-Level Determination

Even though hylomorphism implies a commitment to necessitation and supervenience, it does not imply a commitment to *lower-level determination*, the claim that all of a thing's properties are determined by the properties of its parts and surrounding materials. Determination is a type of necessitation, but it implies more than that; it conjoins necessitation with explanation (cf. Kim 2005: 17ff.):

> *Determination*: Necessarily, for any *x*, if *x* has an *A*-property, *F*, then there is a *B*-property, *G*, such that *x* has *G*, and *x*'s having *G* explains *x*'s having *F*, and necessarily, for any *z*, if *z* has *F*, then *z* has *G*, and *z*'s having *G* explains *z*'s having *F*.

According to this definition, *A*-properties not only covary with *B*-properties, but something's *B*-properties explain why it has the *A*-properties it has. Lower-level determination in particular claims that the activities of a whole are both necessitated and explained by the activities of its parts and surrounding materials:

> *Lower-level determination*: Necessarily, for any structured individual, *x*, if *x* engages in activity *A* at time *t*, then (a) *x* has proper parts and surrounding materials *with powers, P_1, P_2, . . . , P_m, such that the manifestations of P_1, P_2, . . . , P_m at t explain x's A-ing at t*, and (b) necessarily, for any individual *z* and time *t**, if *z* has proper parts and surrounding materials at *t** exactly similar to *x*'s at *t*, *with powers, Q_1, Q_2, . . . , Q_m, exactly similar to the powers of x's proper parts and surrounding materials*, then *z* engages in *A*-ing at *t**, *and the manifestations of Q_1, Q_2, . . . , Q_m at t* explain z's A-ing at t**.

This definition is the same as the definition of hylomorphic necessitation, except for the addition of the italic clauses, which express explanatory relations. If lower-level determination is true, all the activities of a structured whole are explained by the activities of its parts and surrounding materials.

Lower-level determination should not be confused with what we might call *Laplacean determination*, the kind of determination relation that motivates the problem of free will and determinism:

> *Laplacean determination*: Necessarily, for any total state, *S*, of the universe at time t_1, there is exactly one possible resultant state, S^*, that comes about at time t_2 because of *S* in conjunction with the laws of nature.

Unlike lower-level determination, Laplacean determination is a diachronic relation: it claims that a total state of the universe at a time necessitates and explains the total state of the universe at a subsequent time. Moreover, unlike lower-level determination, Laplacean determination is not a relation between the activities of a structured whole, on the one hand, and the activities of lower-level materials, on the other. Laplacean determination has no bearing on the present discussion.

Lower-level determination does have a bearing, however. The reason is that there is a worry that it might undermine the explanatory status of higher-level conditions. If all of Eleanor's activities are explained by the properties of her parts and surrounding materials, then there is little or no explanatory work to be done by any properties she possesses as a composite whole, such as her thoughts and feelings. Consequently, if hylomorphism were committed to lower-level determination, the explanatory status of higher-level properties would be called into question. But hylomorphism rejects lower-level determination.

It might be plausible to endorse something like lower-level determination when it comes to some properties of structured wholes, in particular, their structure-independent properties. Aggregative properties such as mass are examples (Wimsatt 1985). It is plausible to suppose that the mass of a structured whole is determined by the masses of its fundamental physical constituents since those constituents have a collective mass irrespective of how they are spatially arranged. But hylomorphists deny that all the properties of structured wholes are like this. Some properties are structure-dependent or emergent: they depend not just on the materials composing a thing, but on the way those materials are structured or organized (Section 6.3). When it comes to these structure-dependent properties, say hylomorphists, lower-level determination is false. The higher-level activities of a structured whole might covary in all metaphysically possible worlds with certain lower-level material conditions, but those lower-level conditions do not explain why the structured whole engages in the higher-level activities it does, at least not in the way exponents of lower-level determination envision. The reason is that the manifestations of the powers of a structured individual's parts and surrounding materials at a time do not by themselves compose the activities in which a structured individual engages. The manifestations of the powers of those parts and surrounding materials must be structured in the right way as well.

Claiming that lower-level conditions are sufficient to explain the higher-level activities of a structured whole is tantamount to denying that *structure makes a difference*; it is tantamount to denying that structure is a basic explanatory principle. Since structure does make a difference on the hylomorphic view, knowing only that a's parts and surrounding materials are manifesting their powers, $P_1, P_2, \ldots, P_m$, at time t still does not explain why a is engaging in A-ing at t, for engaging in A-ing requires that the manifestations of $P_1, P_2, \ldots, P_m$ be structured A-wise. The A-wise structure is explained not by the activities of a's parts, but by a itself—its manifesting its power to impose an A-wise structure on the way its parts and surrounding

materials manifest their powers. This is a power that *a*'s parts by themselves do not have. The reason is that taken by themselves, those parts do not have the individual-making structure that makes *a* a unified whole. It is that structure which confers on *a* the power that its parts lack. *a*'s distinctive powers and the activities in which they enable *a* to engage correspond to distinctive kinds of descriptions and explanations. Those descriptions and explanations typically appeal to the higher-level powers and activities of *a* rather than to the powers and activities of its parts and surrounding materials. If, for instance, we want to know why Madeleine is playing the piano in contrast to doing her homework, the explanation will typically appeal to things like Madeleine's thoughts and feelings. A description of her neuromuscular systems will not supply a relevant answer to the why-question, it will fail to explain her actions.

The foregoing doesn't imply that the powers and activities of Madeleine's parts and surrounding materials are irrelevant to a complete explanation of what she does. Her behavior, after all, depends on those parts and materials for reasons discussed in Section 8.2: the activities of those parts and materials compose her activity. Moreover, there might be cases in which we seek an explanation that does appeal to Madeleine's parts and surrounding materials—cases, for instance, in which we want to know about the kinds of physiological subsystems that enable Madeleine to do the things she does. To appreciate these points, it will be helpful to discuss different kinds of explanation.

10.3 Causal-Explanatory Pluralism

Explanations are answers to certain kinds of questions. Bas van Fraassen has argued that they are answers to why-questions (1980: 134). Elsewhere, I've argued that some explanations are answers to how-questions as well (Jaworski 2009). If explanations are answers to why- and how-questions, then the logic of those questions can provide a starting point for cataloging different kinds of explanations and the factors and relations they cite. The reason is that a question contains the logical form of its possible answers, a principle of erotetic logic known as Hamblin's dictum: "knowing what counts as an answer is equivalent to knowing the question" (Belnap and Steel 1976: 35). Studying a question's logic can reveal the range of possible answers to it. Consequently, if explanations are answers to certain kinds of why- and how-questions, then studying the logic of those questions can reveal the range of things that count as explanations, and knowing what kinds of things count as explanations can reveal in turn the kinds of explanatory factors and relations that exist in the world.

When it comes to why-questions, van Fraassen (1980) has argued that they have three components: a topic, a contrast class, and a relevance relation. He illustrates this with an example: the sentence 'Why did Adam eat the apple?' could express any of the following questions:

Q1. Why did Adam eat *the apple* (in contrast to having eaten something else)?

Q2. Why did Adam *eat* the apple (in contrast to having done something else with it)?
Q3. Why did *Adam* eat the apple (in contrast to someone else having eaten it)?

This illustrates that explanation is contrastive: a why-question always presupposes a contrast class of propositions. In the case of Q1, the contrast class consists of the propositions 'Adam ate the apple,' 'Adam ate the pear,' 'Adam ate the mango,' 'Adam ate the banana,' 'Adam ate the strawberry,' and so on. One of these statements, 'Adam ate the apple,' is what van Fraassen calls the question's *topic*. When we ask a why-question, we assume that its topic is true, and we assume that the other statements in the contrast class are false. In addition, every why-question assumes a relevance relation to the topic and contrast class. Not just any true statement regarding them can count as an answer. To illustrate this, imagine that you and I are roommates, and that one day you come home to discover a hideous and disturbingly erotic sculpture on the coffee table in our living room—a space we often use to entertain guests, including your very staid friends and co-workers. The doubts you have long harbored about my aesthetic sensibilities now come to fore: "Why," you ask, "is that object on our coffee table (in contrast to somewhere else)?" I respond: "It is on our coffee table (in contrast to somewhere else) because its atoms are on our coffee table (in contrast to somewhere else)." It should be obvious that this response does not answer your question. Although it addresses the question's topic and contrast class, it does not do so in a way that is *relevant* to your interests. The causal factors that you are interested in are my *reasons* for choosing to place the object on our coffee table as opposed to somewhere else. We can nevertheless imagine a context in which my answer would be relevant: you and I are debating the merits and demerits of various theories of material constitution. You want to know what it is on my theory that explains the location of composite objects such as the object on our coffee table. You ask: "Why is that object on our coffee table (in contrast to somewhere else)?" "Because its atoms are there," I respond. In this case, my answer is entirely relevant to your question. Different why-questions can thus be distinguished on the basis of their topics, contrast classes, and relevance relations.

There are many different kinds of how-questions as well (Jaworski 2009, 2011: 290–6). Some examples illustrate this:

Swat team captain: "Headquarters, there's a high-yield explosive timed to detonate in four minutes! How do we disarm it?"
Headquarters: "Very carefully!"

The answer strikes us as ridiculous because we assume the captain was requesting a method for disarming the bomb—a series of steps, not a directive of manner. Likewise,

"How did Judith kill Holofernes?"
Answer A: "With a mixture of revulsion and determination."
Answer B: "With a mixture of bile and snake venom."
Answer C: "With a mixture of seduction and cunning."

The first answer supplies the manner in which Judith killed Holofernes; the second supplies the method; and the third supplies the means. Manners, methods, and means are among the things we ask about using how-questions. We also ask for cognitive resolution when the context implies a set of claims which seems jointly improbable, perhaps inconsistent, and a request is made to remove the appearance of improbability. The familiar how-questions that express philosophical problems are examples of these requests:

How can we be free and yet live in a deterministic universe?
How can God exist and yet evil exist as well?

By contrast with how-questions of cognitive resolution, analytic how-questions request a description of steps which contribute to the accomplishment of some activity or procedure. They include questions that request descriptions of means and methods, and also questions that request descriptions of mechanisms. For instance, in response to the question "How does one dance a swing?" the answer, "First, get a good teacher; then practice, practice, practice," takes the question to ask for means, while the answer, "Backstep, one-two-three, one-two-three," accompanied by a demonstration takes the question to ask for a method. A mechanistic interpretation is harder to imagine, but suppose a student in a neuroscience class complained as follows: "Look, all we've been talking about all semester is how individual neurons perform simple little tasks. What I want to know is how humans manage to perform big complicated tasks. For instance, how does one dance a swing?" The answers, "Usually quite awkwardly," "First, get a good teacher," and "Backstep, one-two-three, one-two-three"—answers which take the question to ask for manner, means, and method, respectively—are clearly out of place. What the student has in mind is an answer such as "The primary motor cortex has a group of cells which generate action potentials in the adjacent cells, and those in turn . . . " Such an answer aims at describing the physiological mechanisms at work in dancing a swing.

How-questions of mechanism are often requests to supply *mechanistic explanations* (Bechtel and Richardson 1993; Glennan 1996, 2002; Machamer et al. 2000; Craver 2007; Bechtel 2008), the kinds of explanations yielded by the method of functional analysis (Sections 6.3–6.4).[1] The aim of functional analysis is to explain how a system is able to perform an activity by describing how its subsystems or parts contribute to that activity. This is what a mechanistic explanation of the system's behavior consists in. We provide a mechanistic explanation of how a human heart or an internal combustion engine operates by describing how the activities of its parts contribute to its activities as a whole.

[1] Chalmers (1996) refers to these as 'reductive explanations.'

Mechanistic explanations are distinct from other explanations of living behavior. This point was made by Plato in a famous passage from the *Phaedo*, in which Socrates describes his experience reading Anaxagoras:

> My wondrous hopes were swept away, my friend, when I proceeded to read, and saw [Anaxagoras] neither appealing to thought, nor citing any of the causes responsible for the ordering of things, but instead citing air, and aether, and water, and many other absurdities as causes. To me it seemed exactly the same as someone saying that Socrates does everything he does with thought, and then in undertaking to state the causes of each thing I do were to say that I am sitting here now because, first, my body is composed of bones and sinews, and the bones are hard and have joints separating them, while the sinews for their part contract and relax, and cover the bones along with the flesh and skin that contains them, and that because the bones move freely in their joints, the contracting and relaxing of the sinews somehow enables me to bend my limbs now, and this is the cause of my sitting here in a bent position . . . But to call such things causes is most absurd. If someone were to say that without having bones, and sinews, and such, I would not be able to do what I believe best, that would be true. But to say that I do what I do *because* of these, and therein act with thought, but not on account of choosing what I believe best—*that* would be an extremely careless way of speaking. (*Phaedo* 98c–99b)

According to Socrates, the problem with Anaxagoras' account of human action is not that Anaxagoras misdescribes the physiological mechanisms at work in human action, but that he assumes those mechanisms are relevant to answering the questions about human behavior Socrates is interested in. Those questions, Socrates thinks, can only be answered by appeal to reasons: thoughts and choices based on beliefs about what is best. Consider another example.

Cecilia is waiting anxiously by the door for Madeleine to come downstairs so they can leave in time for an appointment. Upon being told that Madeleine is reading a book, she asks, "Why is Madeleine reading (in contrast to hurrying down the stairs)?" and receives the following response:

A1. Madeleine is reading (in contrast to hurrying down the stairs) because light reflected off the pages of the book is striking Madeleine's retinas, and the muscles in her eyes are moving in such-and-such ways, and such-and-such neurons are firing in her cortex.

A response along these lines is clearly irrelevant to Cecilia's question. What Cecilia wants is a response that locates Madeleine's behavior within a broader pattern of reasons, as the following response does:

A2. Madeleine is reading (in contrast to hurrying down the stairs) because she thinks finishing the chapter is more important than being on time.

Contrast this example with another: Madeleine's seizures have not responded to drug treatment and the only remaining course of action is lobectomy: doctors must remove part of her brain. To prepare for the operation, they must first identify the

damaged sections of brain tissue. They do so using a new minimally invasive technique: Madeleine wears a cap-like apparatus during her daily activities, which collects data about her brain states for later examination in conjunction with a video record of her activities. When doctors examine the data, they discover an anomaly: the brain regions that in most people are active as they engage in voluntary leg movements, such as hurrying down stairs, are active in Madeleine when she is reading. “Why is Madeleine reading (in contrast to hurrying down the stairs)?” the doctors ask. In this context, they are probably looking for an explanation such as

A3. Madeleine is reading (in contrast to hurrying down the stairs) because during her development certain neural structures had to be “re-wired” to avoid the damaged sections of cortex.

Whatever the details, they are certainly not looking for a response that reveals Madeleine’s reasons for reading; they are not concerned with the rational structure of Madeleine’s behavior, but with something else, namely the states of the various physiological substructures that enable Madeleine to engage in the activities she does.

10.4 Lower-Level Explanations

The logic of why- and how-questions is only a starting point for understanding the range of causes and causal relations there are. The causal factors and relations that exist might outstrip our current ways of trying to conceive of them. There might be causes and causal relations that we do not currently know how to ask about. Likewise, our ways of conceiving those factors might have redundancies that make it seem initially as though we are asking about different factors when in fact we are asking about the same ones. These are the kinds of redundancies that would be revealed if we were successful in reducing one conceptual framework to another. If it turned out, for instance, that thoughts and feelings were identical to states of the nervous system, then the kinds of physiological mechanisms Anaxagoras cites might in principle be able to answer Socrates’ questions, and any dissatisfaction that Socrates and others might feel when given that answer could be attributed not to the answer’s content, but simply to its not being framed in the vocabulary that Socrates and those others prefer.

For reasons that have already been touched upon, hylomorphists are committed to denying that the kinds of questions Socrates asks about human behavior could be answered even in principle by appeal to physiological mechanisms. According to hylomorphists, those questions ask about the rational structure of people’s behavior, and the rational structure of behavior is something different from the activities of the physiological mechanisms that get structured rationally. Hylomorphists are thus committed to denying that explanations like Anaxagoras’ could ever take over the explanatory jobs we expect appeals to reasons to perform. The activities of an individual’s parts and of other lower-level things cannot explain the activities of

the individual as a whole, they say, at least not in the way exponents of lower-level determination intend. If that way of explaining higher-level behavior is supposed to pose a threat to the causal/explanatory autonomy of higher-level activities, then a description of lower-level mechanisms would have to be poised to take over all the explanatory roles (or a significant portion of them) that descriptions of higher-level activities currently play. But if hylomorphists are right, that kind of explanatory takeover is not in the offing. Lower-level mechanisms cannot explain higher-level activities in the sense that generates worries about higher-level causal/explanatory efficacy. Consequently, if this is the sense of 'explain' expressed in the definition of lower-level determination, hylomorphism is committed to rejecting lower-level determination.

This does not mean, however, that lower-level physiological factors cannot explain higher-level behavior in some other sense. Answer A3 in Madeleine's case illustrates this, as does Socrates' remark: "If someone were to say that without having bones, and sinews, and such, I would not be able to do what I believe best, that would be true." An account of lower-level explanation that captures Socrates' idea here can be found in Aristotle's notion of *hypothetical necessity* (*anangkē ex hupotheseōs*):[2]

> [W]hy does a saw have such-and-such features? In order to perform this function and achieve this end. This end cannot come to be, however, if the saw is not made of iron, and so it is necessary for it to be made of iron if the saw is to be and to perform its function. What is necessary is thus conditional . . . if the end is or will be, then the preceding things are or will be as well . . . if the necessary materials do not exist, the end that the process is for will not come to be either . . . Thus, if there is to be a house, it is necessary for these things to come to be or to be present; and in general, the matter that is for something must exist, such as bricks and stones if there is to be a house. It is not because of these that the end comes to be, however, except insofar as they are its matter, nor will it be generated because of them. In general, however, the end (the house or the saw) requires the matter (the stones or the iron). (*Physics* 200a10–29)

What Aristotle says here follows from his more general account of causation and explanation—what is often called the *doctrine of the four causes*. A cause (*aitia*), Aristotle says, answers the question *dia ti*: *Why?* or *On account of what?* (*Physics* 194b16–20). Questions of this sort tend to fall into four categories corresponding to the factors involved in change: (1) the structures of things—their forms, (2) the materials that get structured in those ways—their matter, (3) what is responsible for bringing about the structuring of that matter—what later Aristotelians called the 'efficient cause,' and (4) what that structuring is *for*, what contributions it makes to a yet broader scheme of organization—the end. Answers to these different questions highlight different types of explanatory factors or causes, and various senses of

[2] There are two uses of the term 'necessity,' Aristotle tells us (*On the Parts of the Animals* 639b21–640a10). When it comes to changeable physical things, the operative sense is conditional.

'because' express the various ways that those causes contribute to explaining the occurrence or phenomenon in question.

Living things have distinctive forms that make them the kinds of living things they are. Those forms, however, must be embodied in some suitable matter. What determines the suitability of matter is whether it has the characteristics that enable a living thing to perform its distinctive activities—its function (*ergon*). An eye, for instance, must be made out of materials that are transparent in order to enable seeing (*De Sensu* 438a13–17), just as a saw must be made out of iron if it is to enable sawing. Because living things and their activities depend on material conditions, a complete account of what they are and why and how they can do what they do must include the contributions made both by their form and by their matter. Some aspects of living behavior are explainable by appeal to one but not the other. Material conditions can explain why the eyes are capable of receiving the sensible forms of things, but they do not explain why animals have eyes. The explanation for that appeals instead to the kinds of activities that animals engage in and the way that having eyes contributes to those activities. Similarly, Godehard exists not because fundamental physical particles have various charges or spins, but because the fundamental physical particles composing him are structured human-wise, or (in another sense of 'because') he exists because of the reproductive activity of his parents. Of course, if fundamental physical particles did not have the charges and spins they do, they might be incapable of being structured human-wise. In the sense that there could be no human-wise structuring and hence no Godehard without fundamental physical particles having the properties they have, it is possible to say that Godehard exists (or that Godehard could only exist) because fundamental physical particles have the properties they do. But this is a different sense of 'because' from the ones just considered: it is the sense in which the particles explain Godehard's existence "insofar as they are his matter."

On the hylomorphic view I'm defending, structured individuals and activities are composed of lower-level materials or activities that have the right kinds of individual- or activity-making structures. In that sense, structured individuals and activities depend on lower-level things: they cannot exist without them. We've seen that the essential embodiment of hylomorphic structure implies a commitment to hylomorphic necessitation and supervenience: higher-level phenomena are both necessitated by and supervenient upon lower-level phenomena (Sections 9.1–9.2). But hylomorphism denies that higher-level phenomena are determined by lower-level phenomena in the sense expressed by lower-level determination. The latter conjoins necessitation with explanation. It is true only if lower-level phenomena not only necessitate higher-level phenomena, but also explain them. I've argued, however, that hylomorphism denies that lower-level phenomena can explain higher-level phenomena in the sense that is relevant to lower-level determination. In that sense of 'explain,' lower-level factors are supposed to pose a potential threat to the efficacy or the autonomy of higher-level explanatory factors. According to lower-level determination, if all of an individual's activities are explained by lower-level factors, then

there is little or no explanatory contribution for higher-level factors to make. The hylomorphic view I've described denies that lower-level factors explain things in this way. Lower-level factors are just some of the many factors that contribute to something's behavior—just some factors of the many that belong to its 'complete cause' in Mill's (1843) sense. Those factors, moreover, contribute in different ways which correspond to different senses of 'because.' Eleanor is running because of a desire to stay in shape, and also because of various physiological occurrences. In each of these cases, 'because' means something different; it expresses a different kind of contribution to Eleanor's behavior. One contributes as the formal, efficient, or final cause, the other contributes as the matter. In different contexts, we select which kind of factor we are interested in and ignore the rest (Section 8.5). Doing so does not amount to denying that those other factors are operative or to denying that they are essential to the phenomenon in question. It merely reflects our limited cognitive capacities. We are only able to understand complex things piecemeal; we cannot take in their complete causes all at once.

10.5 Objections

There are at least two objections to the line of reasoning defended in this chapter. The first argues that structo-physical necessitation implies that lower-level physical conditions can explain higher-level powers even if they cannot explain higher-level activities. According to the hylomorphic theory I've defended, says the objection, the powers of a structured individual are essentially embodied in its parts, but structo-physical necessitation seems to imply that lower-level conditions are able to explain why a structured individual has the parts it has. Suppose again that the *p*s compose Godehard, and that some physical event, *e*, results in the *q*s becoming spatially arranged in exactly the way the *p*s are. In that case, it follows from structo-physical necessitation that the *q*s come to compose an exact replica of Godehard, namely Godehard*, and his biofunctional parts. Since those parts embody all of Godehard*'s powers, if we ask why he has the powers he has, the answer seems to be that he has them because of *e*, because that event arranged the *q*s spatially in such a way that they came to compose him and all the parts that embody his powers. A physical event is thus able to explain the powers of a structured individual. Consequently, says the objection, even if a thesis like lower-level determination does not apply to the activities of structured individuals, it might still apply to their powers.

The objection, however, is flawed. It makes out that the explanatory situation is much simpler than it is in fact. To see this, contrast the following questions:

Q1. How did the *q*s come to be arranged spatially at *t* in a way exactly similar to the way the *ps* are arranged (in contrast to being arranged spatially in some other way)?
Q2. Why are the *q*s structured human-wise at *t* (in contrast to being structured in some other way or in no way at all)?

Q3. How did Godehard* come into existence at *t* (in contrast to not having come into existence at t)?
Q4. Why does Godehard* have the powers he has at *t* (in contrast to having other powers)?
Q5. Why does Godehard* have the biofunctional parts he has at *t* (in contrast to having different biofunctional parts)?

On the hylomorphic view, the occurrence of event *e* provides an answer only to Q1: it explains why the *q*s came to be spatially arranged in exactly the way the *p*s are. It does not answer Q2. On the hylomorphic view, what explains why the *q*s are structured human-wise at *t* (in contrast to otherwise) is that Godehard* is configuring them human-wise (in contrast to otherwise) at that time. It is true that event *e* spatially arranged the *q*s like the *ps*, and necessarily, on the hylomorphic view, if physical materials are arranged in that way, then those materials must compose a human or humanlike individual. But this merely states a correlation between physical materials being spatially arranged in a certain way and those materials composing a structured individual. Composition involves more than a spatial arrangement of physical materials on the hylomorphic view: it involves structure. The *q*s compose Godehard* at *t* because Godehard* imposes a human-wise structure on them. If structo-physical necessitation is true, then it is impossible for there not to be a structured individual who imposes on the *q*s a human-wise structure if the *q*s are arranged spatially at *t* in exactly the way the *p*s are. But this does not imply that Godehard* imposes a human-wise structure on the *q*s *because* they are spatially arranged in the way they are. He imposes that structure on the *q*s because he is a structured individual and configuring materials is what structured individuals essentially do. This is what it means to say that structure is a *basic* ontological principle. It provides a terminal explanation for why some things are the ways they are.

For similar reasons, the occurrence of *e* does not answer Q3. It explains how the *q*s came to be arranged in the way necessary for a structured individual like Godehard* to exist, but there is more to the explanation of how Godehard* came into existence than that. His existence depends in part on hylomorphic structure. If there were no hylomorphic structures, then no matter how the *q*s were spatially arranged, they would compose nothing on the hylomorphic view. Structure is thus an essential component of the answer to Q3, a component that event *e* does not supply.

Something analogous is true of Godehard*'s powers and parts, which is why event *e* doesn't answer questions Q4 and Q5 either. The reason Godehard* has the powers he has is that he has a human-wise structure. It is that structure that confers on him his distinctively human powers. This is what it means to say that *structure makes a difference*; structure explains why structured individuals have the distinctive powers they do. Likewise, on the hylomorphic view, something's status as a part depends on its contributing to the activities of a whole. The kind of part it is, moreover, depends on the kind of contribution it makes: different parts contribute to something's

activities in different kinds of ways (Sections 6.3–6.4). An explanation of why Godehard* has the biofunctional parts he has thus depends on the kinds of activities in which Godehard* engages, and activities, on the hylomorphic view, depend on structure. Godehard* has the biofunctional parts he has because humanlike individuals engage in activities, and those activities are composed of the subactivities of very specific kinds of parts. The objection is thus based on a misunderstanding of the explanatory situation.

A second objection to the line of reasoning defended in this chapter comes from someone inspired by David Chalmers (1996: 48–50). Chalmers claims that supervenience implies a commitment to reductive explanation. Someone might run with this claim and argue that if Chalmers is right, then hylomorphists must be committed to lower-level determination after all since they are committed to necessitation, and this implies strong supervenience, and strong supervenience implies in turn the possibility of reductively explaining higher-level phenomena.

There are two things to say in response to this objection. First, by 'reductive explanation,' Chalmers means roughly the kind of explanation I have been referring to as 'mechanistic explanation,' an explanation that answers a how-question of mechanism by appeal to functional analysis (although see Section 12.5 for an important caveat). But mechanistic explanation cannot be the kind of explanation that is operative in the definition of lower-level determination. The reason is that mechanistic explanation does not pose a threat to the explanatory status of higher-level conditions. If there are mechanistic explanations available for all of Eleanor's activities, it does not follow that there is little or no explanatory work to be done by any higher-level properties she has since those properties might answer requests for different kinds of explanatory information. The notion of explanation that figures in the definition of lower-level determination must pose a credible threat to the explanatory status of higher-level properties: it must imply that there is no causal or explanatory work left for higher-level properties to perform once mechanistic explanations have done their bit. This is not the notion of explanation Chalmers refers to as 'reductive explanation.' He himself admits that "reductive explanation is not the be-all and end-all of explanation" (1996: 49); it does not exhaust all the explanatory jobs there are to perform. This by itself is enough to dispense with the objection. But there is more to say.

There is some reason to think that Chalmers is wrong about supervenience implying a commitment to reductive explanation. To appreciate this, let us start with the definition of supervenience Chalmers works with—what he calls 'global logical supervenience':

> [F]or any logically possible world *W* that is [B]-indiscernible from our world, then the [A]-facts of our world are true of *W*. (1996: 40)[3]

[3] Note that Chalmers uses 'B' to refer to the supervening properties and 'A' to refer to the subvening ones. I've followed Kim, McLaughlin, and others in doing the reverse. To avoid unnecessary incongruence

By 'facts,' Chalmers appears to have in mind the kinds of things that we have been calling 'events' or 'states of affairs': individuals having properties or standing in relations at times. He says, for instance, "I use 'A-fact' as shorthand for 'instantiation of an A-property'", although he adds, "the identity of the individual that instantiates an A-property is irrelevant to an A-fact as I am construing it; all that matters is the instantiation of the property" (1996: 361 n. 2). This addition suggests that he takes facts to be like events in Chisholm's (1970) sense rather than Kim's (1973, 1976): they are more like universals than particulars. This point will become relevant momentarily.

The first thing to recognize about this definition of supervenience is that it does not connect in any way to Chalmers' notion of reductive explanation. Reductive explanations in his sense involve functional analyses of complex activities into simpler subactivities performed by lower-level items such an individual's parts and surrounding materials. Yet, even if mental properties supervene on physical properties, it does not follow from Chalmers' definition of supervenience that mental properties can be functionally analyzed and that there must be lower-level candidates to perform the functions those analyses postulate. To secure this result, we need to revise Chalmers' definition along the following lines:

> [GS] *A*-properties supervene on *B*-properties $=_{\text{df.}}$ for any world w, if the *B*-properties of lower-level individuals $b_1, b_2, \ldots, b_n$ in w are indiscernible from the *B*-properties of $b_1, b_2, \ldots, b_n$ (or their proxies) in the actual world, then the *A*-properties of the individuals composed of the *b*s (or their proxies) in w are indiscernible from the *A*-properties of the individuals (or their proxies) composed of the *b*s (or their proxies) in the actual world.

GS defines a supervenience relation for multiple domains: a domain comprising composite individuals and their properties, on the one hand, and a domain comprising the materials composing those individuals and the properties of those materials, on the other. The addition of the expression 'or their proxies' in GS is meant to accommodate Chalmers' conception of facts. For Chalmers, the identity of facts does not depend on the individuals instantiating properties. If properties $A_1, A_2, \ldots, A_n$ are instantiated by individuals $a_1, a_2, \ldots, a_n$, respectively, in world w, and those same properties are instantiated by distinct individuals $b_1, b_2, \ldots, b_n$ in w^*, then worlds w and w^* could still qualify as *A*-indiscernible by Chalmers' lights. To accommodate this possibility, we need to capture the idea that supervenience does not require that numerically the same individuals that exist in w also exist in w^*; it is sufficient for suitable proxies to exist in that world. What are proxies? Intuitively, they are individuals that play in one world the same instantiating roles that different individuals play in some other world. Paull and Sider (1992) define a similar notion

with the discussion of supervenience in this chapter, I've switched the 'A' and the 'B' in the quotation from Chalmers.

in terms of a proxy-identifying bijection that maps individuals from one world onto individuals in another. Let us assume that some such function is available and that the notion of a proxy individual is a workable one.

Even if we assume all of the foregoing, it still does not follow that a commitment to supervenience implies a commitment to reductive explanation, at least not for mental properties. To appreciate this, let us suppose that mental properties supervene on physical properties in the way GS specifies:

> [MPS] For any world *w*, if the physical properties of lower-level physical materials b_1, $b_2, \ldots, b_n$ in *w* are indiscernible from the physical properties of $b_1, b_2, \ldots, b_n$ (or their proxies) in the actual world, then the mental properties of the individuals composed of the *b*s in *w* are indiscernible from the mental properties of the individuals (or their proxies) composed of the *b*s (or their proxies) in the actual world.

Suppose now that substance dualism were true as a matter of logical necessity, that in every possible world if there were individuals that have mental properties, those individuals would not have any physical properties, nor would they be composed of any physical materials. In that case, it would follow that any individuals composed of physical materials in any possible worlds would have the same mental properties, namely none. They would thus be mentally indistinguishable. Consequently, any object composed of the *b*s in *w* would be mentally indistinguishable from any object composed of the *b*s (or their proxies) in the actual world. As a result, if the substance dualistic theory I've described were true, the consequent of MPS would also be true, for there would be no individuals with mental properties in any possible world *w* or in the actual world that were composed of physical materials, and we've seen that from this it follows trivially that the mental properties of every individual that is composed of physical materials in one of those worlds would be indiscernible from the mental properties of every individual composed of those materials (or their proxies) in any world. But if the consequent of MPS is true, then MPS as a whole is true; mental properties supervene on physical properties. The assumption that substance dualism is necessarily true is thus compatible with the assumption that MPS is true: it is compatible, in other words, with the assumption that mental properties supervene on physical properties in all possible worlds; in fact, it actually entails supervenience in those worlds. Yet, this does not imply that mental properties are reductively explainable by appeal to physical properties, for if substance dualism is necessarily true, then mental properties will not be functionally analyzable in any possible world into subactivities performed by lower-level materials, for in any world in which substance dualism is true, mental properties will not be physically realized or embodied, and in that case, there will be no analysis of those properties into the subactivities of realizing or embodying subsystems.

Based on the foregoing considerations, we can conclude that the ideas Chalmers advances do not establish that a commitment to supervenience implies a

commitment to reductive explanation.[4] Hylomorphists are free to endorse necessitation and supervenience while yet denying lower-level determination.

10.6 Williams' Worry

Despite what's been said about the hylomorphic rejection of lower-level determination, and the irreducibility of structural discourse, someone might still harbor suspicions that hylomorphism is just a form of physicalism in disguise. As Bernard Williams once put it, "hylomorphism emerges as just a polite form of materialism" (1986: 224). I will call the suspicion that hylomorphism is really just a form of physicalism *Williams' worry*. There are a number of forms Williams' worry could take; some concern physicalism, others reductivism. Here are a few:

Worry 1: Biological and physical structure

Hylomorphism's central notion is structure. But the kinds of structures hylomorphists have in mind—biological structures, for instance—are just complex relations among physical particles, and if that's the case, then surely those structures can be exhaustively described and explained by physics. Hylomorphism thus appears to be a kind of physicalism.

Worry 2: Exhaustive decomposition

Hylomorphism is committed to structured wholes being exhaustively decomposable into fundamental physical materials. Likewise, it is committed to the activities of structured wholes being decomposable into the activities and subactivities of their parts and surrounding materials—which are all exhaustively decomposable into fundamental physical materials. Surely the exhaustive decomposition of structured wholes and their activities into lower-level individuals and their activities is tantamount to some kind of physicalism. Hylomorphism is thus a physicalist theory.

Worry 3: Necessitation and physicalism

Hylomorphism implies a commitment to necessitation and strong supervenience. Yet, a commitment to necessitation and/or strong supervenience is a sufficient condition for physicalism; in fact, many philosophers define physicalism by appeal to these notions (Lewis 1983b; Stoljar 2010). Hylomorphism is thus a species of physicalism.

Worry 4: Necessitation and reduction

Hylomorphism is committed to necessitation and strong supervenience. Yet, there are good reasons to think that necessitation and/or strong supervenience implies some

[4] In fact, Chalmers (1996: 48) himself appreciates that the sufficiency of supervenience for reductive explanation is much less compelling than the necessity of supervenience for reductive explanation.

type of reductivism. This premise was the basis of one of Kim's (1984, 1990) challenges to the coherence of nonreductive physicalism. Many nonreductive physicalist theories imply necessitation/supervenience, but if necessitation/supervenience implies reductivism, then those theories cannot claim to be nonreductive. If we endorse Kim's premise, then we have to draw a similar conclusion about hylomorphism: contrary to what's been said, hylomorphism is committed to the idea that descriptions and explanations that appeal to thoughts, feelings, perceptions, and intentional actions are reducible to descriptions and explanations that appeal to physiological mechanisms alone. Hylomorphism is thus committed to reductivism.

Worry 5: Phenomenal consciousness

According to hylomorphism, thought, feeling, and perception are composed of the structured manifestations of the powers of lower-level things. In the process of describing hylomorphism's approach to these phenomena, I've made no mention of phenomenal consciousness. On the contrary, I've said that the lower-level individuals, powers, and manifestations that compose thought, feeling, and perception are revealed through functional analysis. This suggests a view according to which phenomenal consciousness is a physical process. If 'consciousness' refers to anything in a hylomorphic framework, it must refer to something that can be understood in physical terms. Hylomorphism thus seems committed to some type of physicalism.

Worry 6: Epiphenomenalism and lower-level laws

Hylomorphism claims that the behavior of structured individuals like us never violates the laws governing their fundamental physical constituents, that it is only because lower-level materials retain their distinctive powers that structured individuals can recruit them as components that contribute to their own activities. But if structured individuals can never violate fundamental physical laws, then it seems that it must be impossible for our thoughts and feelings to make any causal difference to our behavior. How, after all, could those thoughts and feelings make a difference, if not by intervening at the level of fundamental physics? Hylomorphism must be committed, therefore, to something like epiphenomenalism. Even if there are structures that elude physical description, the things we do, our behavior, can still be exhaustively explained by appeal to physics alone.

Worry 7: Epiphenomenalism and Laplacean determination

There is another way of bringing out the same worry. It concerns the notion of Laplacean determination described in Section 10.2:

> *Laplacean determination*: Necessarily, for any total state, *S*, of the universe at time t_1, there is exactly one possible resultant state, S^*, that comes about at time t_2 because of *S* in conjunction with the laws of nature.

If Laplacean determination is true, then it seems that whatever states of affairs come about in the universe come about because of laws of nature and antecedent conditions. But if that is true, then it seems that higher-level structural phenomena have no explanatory role to play in the world, they are mere epiphenomena: everything that happens can be exhaustively explained by appeal to physics alone. Hylomorphism is thus committed to physicalism.

Worry 8: Behaviorism and the psychophysical identity theory

Hylomorphism seems to be committed to some type of reductive physicalism, either behaviorism or the psychophysical identity theory. Hylomorphism claims that thoughts, feelings, and perceptions are coordinated manifestations of the powers of our parts and surrounding materials. It claims, moreover, that we can directly perceive at least some of these manifestations in our pedestrian dealings when other people think, feel, or perceive, and that because of this there is no problem of other minds (Section 13.5). The idea that we can directly perceive other people's thoughts, feelings, and perceptions, however, sounds like behaviorism. Hylomorphists might reply that there is more to a thought, feeling, or perception than the coordinated manifestation of powers we can directly perceive. There are, for instance, parts of the nervous system that manifest their powers when other people think, feel, or perceive but that are not directly observable in our pedestrian dealings. These are among the parts that manifest their powers in a coordinated way when we remember, imagine, fantasize, or think "in our heads," without revealing what we are remembering, imagining, fantasizing, or thinking about. Yet, the idea that our private remembrances, imaginings, fantasies, and thoughts are coordinated manifestations of the powers of parts of our nervous systems sounds like a kind of psychophysical identity theory. Hylomorphism thus appears to be committed to something like behaviorism when it comes to some mental states, and it appears to be committed to something like the psychophysical identity theory when it comes to others. Either way, it appears to be committed to something like reductive physicalism because both behaviorism and the psychophysical identity theory imply that physical discourse can in principle take over the descriptive and explanatory roles that psychological discourse plays.

Worry 9: Multiple realizability

Hylomorphism seems incompatible with a central tenet of the antireductionist consensus in the philosophy of mind, namely multiple realizability, the claim that higher-level activities and capacities, such as psychological ones, can be correlated with lower-level material conditions of many different kinds. Even though thought, feeling, perception, and intentional action are correlated with states of a human nervous system in you and me, they could be correlated with states of silicon circuitry in sophisticated robotic systems or with nonhuman physiological states in members of alien species. According to hylomorphists, however, our thoughts, feelings,

perceptions, and actions are essentially embodied in the kinds of parts we possess. The activities of those parts and surrounding materials constitute material conditions that, according to hylomorphists, are metaphysically necessary for us to engage in higher-level activities. Hylomorphists, moreover, appear to construe those material conditions so narrowly that they rule out the standard examples of multiple realizability. For example, hylomorphists would deny that a human could have all of its parts replaced by nonbiological prostheses. Moreover, since the activities in which a human engages are composed partly by the activities of its parts, a human who underwent a replacement of a substantial number of its biological parts with nonbiological prostheses would no longer be capable of engaging in a range of its typical activities—including perhaps thinking, feeling, perceiving, and acting. Given reasonable assumptions, this appears to imply that a robotic being could never think, feel, and perceive. Moreover, if the predicates and terms we use to refer to our thoughts, feelings, perceptions, and actions refer to activities that are essentially embodied in the kinds of parts we possess, then there is reason to think that according to hylomorphism, members of alien species that differ from us physiologically could never think, feel, perceive, or act—in fact, hylomorphism may even imply that familiar nonhuman animals such as dogs and cats cannot think, feel, perceive, or act. Some of these implications are implausible in their own right, and they also contradict the multiple realizability thesis. That thesis, however, has been a mainstay of antireductionist consensus since the early 1970s. Because hylomorphism appears committed to rejecting it, it seems that hylomorphism must be committed to some type of reductivism.

There might be other forms that Williams' worry could take, but these nine are a good starting point. What I plan to argue in the chapters that follow is that each worry is based on at least one false premise. Demonstrating this will require a preliminary discussion of physicalism and reductionism since I think confusions about exactly what these entail are largely responsible for motivating the worries. Clearing up those confusions will be the goal of Chapter 11. I will then return to Williams' worry in Chapters 12 and 13. I address Worries 1–4 in Sections 12.2–12.4; Worry 5 in Sections 12.5–12.7; Worries 6 and 7 in Section 13.4; Worry 8 in Section 13.6, and Worry 9 in Section 13.7.

10.7 Conclusion

One of the most common misunderstandings about the nature of necessitation and supervenience relations is that they entail some type of explanatory condition. On this misguided view, saying that A-properties supervene on or are necessitated by B-properties implies that B-properties explain A-properties. But it can be shown in several ways that necessitation and supervenience do not entail any such explanatory condition: each merely states a pattern of property covariation; neither purports to explain why that pattern obtains, and neither implies that lower-level conditions

explain higher-level ones. This point played a crucial role in criticisms of nonreductive physicalism. It is also what distinguishes necessitation and supervenience from determination.

Determination conjoins necessitation with explanation. It claims not merely that *A*-properties covary with *B*-properties, but that something's *B*-properties explain why it has the *A*-properties it has. Lower-level determination, in particular, is the thesis that all of a thing's properties are determined by the properties of its parts and surrounding materials. Lower-level determination is associated with worries that higher-level properties, states, or conditions might have no causal or explanatory import. If all of Eleanor's activities are explained by the properties of her parts and surrounding materials, then there is little or no explanatory work to be done by any properties she possesses as a composite whole. Hylomorphism rejects lower-level determination because lower-level determination is incompatible with the claim that structure is a basic explanatory principle. Some properties are structure-dependent or emergent: they depend not just on the lower-level materials composing or surrounding a thing, but on the way those materials are structured or organized. If there are structure-dependent properties, then lower-level determination is false. Even if the higher-level activities of a structured whole covary in all metaphysically possible worlds with certain lower-level material conditions, it still does not follow that those lower-level conditions explain why the structured whole engages in the higher-level activities it does—at least not in a sense of 'explain' that poses a threat to the efficacy or autonomy of higher-level explanations. Some of an individual's behavior can be explained only by appeal to higher-level structures. Appeals to higher-level structures and appeals to lower-level conditions answer requests for different kinds of information—information about structure, on the one hand, and information about lower-level things that are structured, on the other. On the hylomorphic view, both structure and things that are structured are equally real, autonomous, and efficacious explanatory factors or causes.

The hylomorphic account of thought, feeling, perception, and intentional action can engender what I call *Williams' worry*, the worry that hylomorphism is, as Bernard Williams put it, "just a polite form of materialism." Williams' worry can take a number of forms, which concern, among other things, hylomorphism's account of composition, its account of necessitation and supervenience, its approach to phenomenal consciousness, multiple realizability, and downward causation. In order to address these worries adequately, it is necessary to get clear on the nature of physicalism and reductivism. This is the goal of Chapter 11.

11

Physicalism and Other Mind-Body Theories

11.1 General Definitions of Mind-Body Theories

The goal of this chapter is to clarify the nature of physicalism and reductivism in order to address Williams' worry, the suspicion that hylomorphism is really just a physicalist theory in disguise. That suspicion, I contend, is based on misunderstandings about what physicalism and reductivism are—misunderstandings that are abetted by the terminology many philosophers have used to discuss them. To avoid confusion, it will be helpful to situate physicalism and reductivism within the broader range of mind-body theories.

Mind-body theories can be categorized in terms of the kinds of properties they countenance. Properties, recall, are supposed to be the nonlinguistic correlates of at least some predicates. We saw in Section 2.1 that the latter can be understood as sentence-frames in Strawson's (1974: 37–8) sense, linguistic expressions such as '___ is red' and '___ is taller than ___' that form sentences when the blanks are filled in by terms. (For convenience, I will continue to omit the blanks when referring to predicates.) Let us agree to categorize properties as mental or physical based on the predicates that express them. Physical properties are properties expressed by physical predicates, and mental properties are properties expressed by mental or psychological predicates. The paradigmatic mental predicates are the predicates of ordinary psychological discourse that we use to express people's propositional attitudes and perceptual or sensory states: 'knows how to solve the problem,' 'believes that it will snow,' 'wants a merit increase,' 'is afraid that the sell order will arrive late,' 'sees the raccoon lumbering across the lawn,' 'feels agitated,' 'has a nagging itch,' and so on. Physical predicates, on the other hand, are the predicates used to formulate theories in physics. How exactly we are to understand 'physics' here is contentious, for among other things, the definition of physics has thorny implications for the definition of physicalism.

Physicalism claims that everything is physical, that everything can be exhaustively described and explained in principle by physics. Physics, however, is progressive. Seventeenth-century physics is not the same as current physics, and current physics is not immune to future revision. This presents physicalists with a dilemma first

described by Carl Hempel (1969). If physics is defined relative to some preliminary stage in its development, a stage at which its theories are still subject to falsification and revision, then physicalism ends up being false since the physical theories advanced at that stage are false; not everything can be exhaustively described and explained by them. If, on the other hand, physics is defined relative to a final, ideal stage of its development, a stage at which physical theory is no longer subject to falsification and revision, then physicalism ends up lacking content. Since we do not yet know what the final, ideal physical theory says, we do not know what it means to say that everything is the way the final, ideal physical theory says it is. The argument thus concludes that physicalism is either false or else lacking content.

Hempel's dilemma is relevant for our purposes because we do not want to begin with a definition of physical predicates and properties that rules out a priori the coherence of physicalism. On this point, we can perhaps do no better than to follow the lead of physicalists themselves. Some physicalists grasp the first horn of Hempel's dilemma: physics should be defined by appeal to current physics (Melnyk 1997; 2003: 14–15). This implies that physicalism is false, but this result is not as bad as one might initially suppose, say these physicalists. If a better future physics replaces current physics, then physicalism will be replaced by a theory that nevertheless closely resembles it and that is a sociological descendent of it. Since this kind of replacement happens unproblematically in physics, we have good reason to think that it can happen unproblematically for physicalism. Other physicalists grasp the second horn of the dilemma (Armstrong 1978a; Lewis 1983b; Poland 1994). They look to define physicalism relative to ideal physics. They concede that this definition lacks some content, but deny that it is completely lacking in content. We might not know exactly what the ideal physical theory will claim, but we know enough to be able to evaluate what physicalism says since past physical theories give us some insight into future ones. If we know something about current physics, this gives us some insight into the possible future physics at which physicists hope eventually to arrive. Physicalists thus have at least two viable strategies for addressing Hempel's dilemma. For our purposes, we do not need to choose between them. We can take 'physics' to refer either to our best current physical theories or to some better future ones. This understanding of 'physics' should provide enough traction to enable us to discuss mind-body theories.

The foregoing definitions of 'mental' and 'physical' should make it evident that being mental and being physical are not mutually exclusive designations.[1] It is

[1] Some philosophers have mistakenly assumed that they are mutually exclusive. Hobbes, for instance, wrongly understood Descartes to endorse the logical exclusion (as opposed to the mere logical independence) of the mental and physical conceptual frameworks in the Third Set of Objections to the Meditations (AT VII, 173). According to Hobbes, Descartes assumes that a thinking thing cannot be corporeal; he assumes that the category of thought *excludes* the category of corporeality. But Hobbes' interpretation is incorrect. In Meditation Two, Descartes considers whether being a body is consistent with his knowledge of himself as a thinking thing, and concludes that it is: "Is it perhaps not truly the case that these very

possible for the same property to be expressed by predicates of both sorts; the property expressed by the predicate 'is in pain' might be the very same property that is expressed by the predicate 'is in brain state B.' The following claim is therefore not trivial:

Property dualism: There are (first-order) mental properties and there are (first-order) physical properties, and (first-order) mental properties and (first-order) physical properties are distinct.

A commitment to property dualism thus defined is what distinguishes monistic theories like physicalism from dualistic ones. The latter accept property dualism, the former reject it.

Among theories that accept property dualism, dual-attribute theories (DATs) claim that individuals with mental properties have some physical properties essentially. Substance dualistic theories deny this. Individuals with mental properties, they say, do not have any physical properties essentially. Consequently, they claim that there are not just two different kinds of properties, there are also two different kinds of substances or individuals: persons, which have only mental properties essentially, and bodies, which have only physical properties essentially.

DATs are sometimes referred to as forms of property dualism, in contrast to forms of substance dualism, but this use of the term 'property dualism' can be misleading since substance dualists are committed to property dualism as well (Armstrong 1968: 11). The label 'dual-aspect' has also been used instead of 'dual-attribute,' but this too can be misleading since 'aspect' suggests that, according to DATs, the mental-physical distinction is merely a matter of how things appear to us. We grasp the same property, albeit in its different aspects depending on the conceptual framework we use to describe it (an idea compatible with the psychophysical identity theory). The term 'attribute' expresses more clearly the ontological nature of the property dualism DATs endorse.

[bodily] things which I am supposing to be nothing, because they are unknown to me, are still in the truth of the matter not different from the 'me' whom I know? *I do not know, and I am not disputing this now*... If this 'I' is taken precisely as such, then it is most certain that knowledge of it does not depend on things the existence of which I do not yet know—not therefore on any of the [bodily] things which I produce in my imagination" (AT VII, 27–8, emphasis added). According to Descartes, knowing that he is a thinking thing leaves his identity with some sort of body an epistemically open possibility; it is consistent with all he knows. He does not think that the application to himself of mental predicates implies that he is not extended and hence not a body; he assumes only that the application to himself of mental predicates does not imply that he *is* a body. He replies to Hobbes: "I did not assume the contrary [i.e. that a thinking thing cannot be something corporeal], nor did I use it in any way as a premise, but left it completely indeterminate until the Sixth Meditation in which it is proved" (AT VII, 175). More recently, John Searle has mischaracterized debates in philosophy of mind in a similar way: "we lack a neutral vocabulary in which to describe mental phenomena... [I]n our traditional vocabulary the most natural characterization is to say that there is a distinction between the mental, on the one hand, and the physical or material, on the other. The mental qua mental is not physical. And the physical qua physical is not mental... This traditional vocabulary assumes the mutual exclusion of the mental and physical; and that assumption creates insoluble problems that have launched a thousand books" (2004: 1–3; cf. 1992: 26).

Monistic theories are of three broad types. Mental monism or idealism claims that everything is mental. Physical monism or physicalism claims that everything is physical, and neutral monism claims that everything is neutral (things are neutral if they can be described and explained by a conceptual framework that is neither mentalistic like ordinary psychological discourse, nor physicalistic like the theories of current or future physics). The remainder of this chapter is concerned with articulating and defending a more precise definition of physicalism.

11.2 The Definition of Physicalism

There are two criteria that any adequate definition of physicalism must satisfy to be applicable to mind-body debates. Tim Crane mentions one of them: "Physicalists might differ in their . . . metaphysical commitments . . . But common to all forms of physicalism is the view that whatever exists is in some sense physical" (1994: 479). Any adequate definition of physicalism must entail the thesis that everything is physical. I'll call this the *core physicalist thesis*.[2] Any theory that is compatible with the existence of nonphysical entities cannot qualify as physicalist. Second, although there are different varieties of physicalism (eliminative, reductive, nonreductive), physicalism is supposed to be a general thesis. A definition of physicalism in general should not imply a commitment to a specific variety of physicalism; it should be compatible with eliminativism, reductivism, and nonreductivism.

Many definitions of physicalism that have appeared in the literature fail to satisfy at least one of these criteria. Examples of definitions that fail to satisfy the first criterion include two mentioned by Peter van Inwagen:

> The thesis that human persons are physical things is called *physicalism*. (This word is also used as a name for the stronger thesis that all individual things are physical things . . .) (2002: 168)

Both of the theses van Inwagen mentions are too weak to satisfy the first criterion of adequacy. Neither implies the core physicalist thesis; both are instead compatible with the existence of nonphysical things. The first definition pertains only to human persons; consequently, it is compatible with the existence of nonphysical things of other sorts such as nonphysical properties or events, or nonphysical individuals that are not human persons. Likewise, the second, stronger thesis van Inwagen mentions pertains only to individuals, so it too is compatible with the existence of nonphysical properties or events. Some DATs are committed to persons being human organisms—we can call them *organismic DATs*. Organismic DATs are compatible with van Inwagen's stronger definition since they claim that human persons are physical individuals, but they also imply that human persons have some nonphysical

[2] In line with what I said in Chapter 2, this should be interpreted as a claim about concrete things. I will put to one side questions concerning abstracta such as numbers, sets, possible worlds, and so on. The core physicalist thesis is not meant to apply to them.

properties, and this is incompatible with the core physicalist thesis. Trenton Merricks offers a definition similar to van Inwagen's, and he is in fact explicit that he does not intend to capture the core physicalist thesis:

'Physicalism' means that human persons are physical and substance dualism is false; it does not mean that everything is physical. (2001b: 183)

Consider also a definition of physicalism offered by John Dupré (who calls it 'materialism'):

[M]aterialism embodies . . . the proposal that everything that happens can be explained, at least in principle, in terms of physical entities and the laws that govern their behavior. (1993: 93)

A crucial ambiguity in Dupré's definition lies in the term 'physical entities.' If physical entities include only those entities postulated by physics, then the definition satisfies the first criterion. Suppose, however, that physical entities include physical individuals, and the latter are defined as individuals that are exhaustively decomposable into fundamental physical materials. The existence of such individuals is compatible with the organismic DAT just described; it is possible for individuals that are exhaustively decomposable into fundamental physical materials to have nonphysical properties, and for their behavior to be governed by laws other than those of physics—emergent laws to match the emergent nonphysical properties. Dupré's definition can thus be understood in a way that fails to satisfy the first criterion.

An example of a definition that fails to satisfy the second criterion is advanced by David Armstrong:

Materialist theories . . . try to reduce mind to body or to some property of body. (1968: 5)

This definition implies that physicalism (what Armstrong calls 'materialism') must be committed to reductivism, a claim that rules out the possibility of eliminative or nonreductive physicalism. Something analogous is true of one of Crane and Mellor's definitions:

[Physicalism] says that mental entities, properties, relations and facts are all really physical. (1990: 394)

This definition rules out eliminativism since it implies that physicalists are committed to there being mental entities; those entities, it says, are identical to physical ones. Consider likewise one of Thomas Nagel's definitions:

Materialism is the view that only the physical world is irreducibly real. (2012: 37)

The terms 'physical world' and 'real' are ambiguous, but on one interpretation, this definition implies that everything is either reducible to physical theory or else unreal; a physicalist, in other words, must be committed either to reductivism or eliminativism. Nonreductive physicalism is left in the lurch.

Nagel offers another definition that seems closer to the mark:

> [E]verything that exists and everything that happens can in principle be explained by the laws that govern the physical universe. (2012: 18)

One possible shortcoming of this definition is that it leaves open the possibility that there might be properties and laws that cannot be described or explained by appeal to physics alone. Consider again an organismic DAT like the one described earlier, which claims that some individuals have emergent properties—nonphysical properties that cannot be described or explained purely by appeal to physics. Suppose that there are law-like connections among these nonphysical emergent properties. These connections would be among the laws that govern the physical universe. Consequently, the resulting view would be compatible with everything that exists and everything that happens being in principle explainable by appeal to the laws governing the physical universe—the emergent laws in conjunction with the laws of physics. Yet, clearly, this view would not be a physicalist one since it would be committed to nonphysical properties and laws. To rule this out, a definition of physicalism should add that the descriptions and explanations provided by physics are exhaustive; they leave out nothing, but provide a complete account of what things are, what they do, how they do it, and why. Another definition offered by Crane and Mellor has a similar shortcoming:

> Physicalists believe that everything is physical: more precisely, all entities, properties, relations, and facts are those which are studied by physics or other physical sciences. (1990: 394)

The organismic DAT described above is compatible with organisms being studied by physics. What it denies is that physics gives us an exhaustive description of the properties organisms have or an exhaustive explanation of the behaviors in which organisms engage.

Consider, by contrast, the following definitions offered by David Armstrong and David Lewis, respectively:

> The world is completely described in terms of (completed) physics . . . [This] is the hypothesis of modern . . . Physicalism. (Armstrong 1978a: 126–7)

> Materialism is the thesis that physics—something not too different from present-day physics, though presumably somewhat improved—is a comprehensive theory of the world, complete as well as correct. The world is as physics says it is, and there's no more to say. (Lewis 1983b: 33–4)

Both of these definitions imply that physics provides a complete or exhaustive inventory of what there is, and in that sense, they mark an advance over the other definitions we've considered. There are nevertheless three amendments I would make to what Armstrong and Lewis say. First, the definition of physicalism should not be confined merely to the descriptive completeness of physics, as Armstrong's definition suggests; it should include the explanatory completeness of physics as well. Second, Armstrong and Lewis both take completed physics (what we earlier called 'final ideal physics') to determine the content of physicalism. There are, however, many

physicalists who would favor a definition in terms of current physics. A definition of physicalism should be general enough to accommodate both ways of defining 'physics.' Finally, both definitions suggest that physics provides a unified descriptive and explanatory framework. Philosophers such as Nancy Cartwright (1999), however, have argued that this suggestion misrepresents the diverse and often disunified conceptual resources of real science. A definition of physicalism should be open to the possibility that philosophers like Cartwright are right. It shouldn't imply a commitment to the unity of physical science, but should be compatible with the idea that our best current or future physics might end up being a patchwork quilt of diverse models. Here is a definition that incorporates these amendments:

> *Physicalism*: Everything can be exhaustively described and explained by the most empirically adequate theories in current or future physics.

This definition appears to imply the core physicalist thesis. The metaphysic developed in Chapters 2–5 implies that the only entities are individuals, properties, and events. If these define the domain over which the definition quantifies, then the definition implies that everything is physical if it implies that all individuals, properties, and events are physical. Consider first, then, whether the definition implies that all properties are physical. To show this, let us assume for *reductio* that physicalism is true at world w and yet not all properties are physical. According to the metaphysic developed in Chapters 2–5, individuals are the primary agents; they are the entities that enter into causal relations, and their properties are the entities in virtue of which they enter into the causal relations they do (Section 2.1). Suppose that F is a property that is not physical. Physical properties are properties that are expressed by the predicates of physics. Since F is not a physical property, it is a property that is not expressed by any of the predicates of physics. Since the predicates of physics are the predicates that factor into the descriptions and explanations physics provides, the kinds of causal relations F enables will not be the kinds of causal relations that physics describes and explains. In that case, however, not everything in w will be describable and explainable by physics, which is contrary to the original assumption that physicalism is true. Consequently, if physicalism is true, all properties must be physical.

Consider now individuals. Individuals must have properties. If physicalism is true, then all properties are physical. Consequently, if physicalism is true, all the properties individuals can have will be physical properties. Since, moreover, properties are the entities in virtue of which individuals enter into causal relations, all the causal relations into which individuals can enter will be causal relations that can be described and explained using only the vocabulary of physics. If physicalism is true, therefore, all the properties an individual has and everything an individual does will be describable and explainable by physics. But surely this is sufficient to qualify an individual as physical if anything is. Consequently, if physicalism is true, all individuals must be physical.

Consider finally events—individuals having properties at times. It seems plausible to assume that if *a*'s having property *P* at time *t* is not a physical event, then either *a* must not be a physical individual or else *P* must not be a physical property. We have just seen, however, that if physicalism is true, all individuals and all properties are physical. Consequently, if physicalism is true, all events must be physical as well. Given reasonable assumptions, therefore, the foregoing definition of physicalism implies the core physicalist thesis; it implies that everything—every individual, property, and event—is physical.

There are at least four objections to the proposed definition of physicalism that are worth considering. The first three objections argue that the definition fails to satisfy the second criterion of adequacy, which says that a definition of physicalism cannot imply a commitment to a specific kind of physicalism such as eliminativism or reductivism. The fourth objection insists that a definition of physicalism must appeal in some way to supervenience or necessitation—something missing from the proposed definition.

The first objection appeals to higher-order properties. At least some higher-order properties are not physical properties, it says, since they are expressed by predicates other than those of physics. It seems plausible to suppose that being in pain, for instance, amounts to being in some physical state that satisfies a corresponding condition *C*. Since 'is in pain' is a predicate not of physics, but of psychological discourse, it seems plausible to suppose that being in pain does not count as a physical property, and the same is true of mental properties in general. Since the proposed definition implies that all properties are physical, it must imply that mental properties do not exist. But that means that the proposed definition fails to satisfy the second criterion for adequacy since it implies a commitment to eliminativism.

This objection depends on the assumption that there are higher-order properties. We saw in Section 5.3, however, that there are good reasons to reject this assumption. According to the Eleatic principle endorsed in Section 2.1, the only existing things are ones that make some sort of causal difference. This implies that the only existing properties are ones that enable individuals to enter into causal relations, but according to the argument advanced in Section 5.3, if there were higher-order properties, they would make no causal difference to their bearers. Consequently, there are good reasons to think that higher-order properties do not exist. There is not, by contrast, good reason to think that mental properties do not exist. Consequently, the objection does not pose a challenge to the proposed definition of physicalism.

The second objection is analogous to the first, but it concerns individuals instead of properties. Limiting the individuals that exist to physical individuals seems unnecessarily restrictive, says the objection. Not all individuals are postulated by physics. In fact, most of the individuals of scientific and pedestrian acquaintance are not postulated by physics: plants, animals, and molecules are examples. Surely a definition of physicalism should not commit us to denying their existence.

This objection assumes that the proposed definition of physicalism is incompatible with the existence of individuals other than those postulated by physics, but this assumption is false. To see this, let us assume that physicalism is true and that a_1, $a_2, \ldots, a_n$ are all the individuals postulated by physics. Assume, moreover, that composition is unrestricted, and that a_j and a_k compose *b*. Since the *a*s are by assumption the only individuals postulated by physics, *b* must not be an individual postulated by physics. Is *b*'s existence compatible with the proposed definition of physicalism? It is difficult to see why it wouldn't be. Someone might argue that it is incompatible on the grounds that the proposed definition implies that *b*'s behavior can be exhaustively described and explained by physics. If everything can be exhaustively described and explained by physics, as the proposed definition implies, then clearly there can be nothing to *b*'s behavior other than what physics can describe and explain. On the basis of this, someone might be tempted to claim further that *b* does not really exist. The argument, however, falls short of establishing this, for it is not clear how *b*'s nonexistence is supposed to follow. Suppose, on the other hand, that composition is restricted, that some *a*s compose something only if they satisfy condition *C*. If physicalism is true, it is plausible to suppose that *C* is a condition that can be exhaustively described and explained by physics. If so, then once again, it follows that there is nothing to *b*'s behavior other than what physics can exhaustively describe and explain since there is nothing to that behavior other than what is contributed by a_j, a_k, their properties, and *C*. Once again, however, what does not follow is that *b* does not exist.

So far, then, there is no reason to think that the existence of composite individuals is incompatible with the proposed definition of physicalism. Of course, not every principle of restricted composition is compatible with physicalism. The kind of restrictivism endorsed by van Inwagen, Merricks, and hylomorphists certainly is not. According to them, composite objects have emergent properties, properties not had by the things that compose them. Consequently, if a_j and a_k compose *b*, then *b* has some properties or engages in some behaviors that cannot be exhaustively described and explained by physics. But aside from this kind of restrictivist view, there does not appear to be any incompatibility between the proposed definition of physicalism and the existence of composite individuals.

A third objection claims that the proposed definition of physicalism implies a commitment to reductivism. We saw a moment ago that the proposed definition implies that there is nothing to the behavior of composite individuals other than what is contributed by their parts, the properties of those parts, and a principle of composition that can be exhaustively described and explained by physics. Surely this amounts to some type of reductivism, says the objection, but according to the second criterion of adequacy, a definition of physicalism cannot imply reductivism. Hence, the proposed definition must be inadequate.

This objection assumes that if something's behavior can be exhaustively described and explained in principle by physics, then this is sufficient for reduction. But it is not. To appreciate this, we will have to consider reduction in greater detail.

11.3 Reduction

The term 'reduction' is used in a variety of ways in philosophy and the sciences. The notion of reduction that interests us here is intertheoretic reduction (Churchland 1986: 278–9). Intertheoretic reduction (simply 'reduction' henceforth) is a synchronic relation between theories or conceptual frameworks in which one of them, the reducing theory or framework, is able to take over the descriptive and explanatory roles played by the other, the reduced theory or framework.[3] Paradigmatically, reduction involves discovering that what we initially took to be two distinct domains are not in fact distinct. Imagine, for instance, that Domain A is included within Domain B, but because of the way we are configured epistemically, we have come to know and describe A-entities in ways different from the ways we have come to know and describe other B-entities. As a result, we describe and explain the behavior of A-entities using a theoretical framework, TA, that is different from the framework we use to describe and explain the behavior of other B-entities, the framework TB. Although we do not initially recognize that Domain A is included in Domain B, we later discover this: we discover that A-entities are really identical to B-entities of a certain sort. Since A-entities are really just B-entities, the principles governing the behavior of A-entities are just special applications of the principles governing the behavior of B-entities in general. Consequently, the behavior of A-entities can be exhaustively described and explained by appeal to TB. Moreover, A-theoretic descriptions and explanations can be systematically reformulated in B-theoretic terms. As a result, TB is able to take over all the descriptive and explanatory roles formerly played by TA. TA is thus reducible to TB. The upshot is ontological and explanatory consolidation or simplification. Where previously there appeared two kinds of entities and two theoretical frameworks for describing and explaining their behavior, there now appears only one. In addition, this consolidation can improve our understanding of A-phenomena. The principles governing A-behavior are shown to be just special applications of B-principles in general, and in that sense, A-principles are explained by B-principles. A-principles obtain to the extent they do because B-principles obtain in A-specific circumstances.

The notion of intertheoretic reduction was given its first rigorous treatment by Ernest Nagel (1961). Although his account of it was shown to have several shortcomings, it still provides a helpful starting point for understanding many of the basic features of reduction. Nagel's account is based on a syntactic model of theories and a derivational model of explanation. A syntactic model of theories claims roughly that theories are sets of law statements, and a derivational model of explanation claims

[3] Nickles (1975) distinguishes intertheoretic reduction in the sense considered here, which he refers to as *domain-combining* reduction, from another kind of intertheoretic reduction, which he refers to as *domain-preserving*. An example of the latter is the case of special relativity and classical mechanics. Special relativity is said to reduce to classical mechanics at low velocities.

that explanation is derivation from law statements. To say that TA is reducible to TB on Nagel's view is to say that the law statements of TA can be derived from the law statements of TB in conjunction with statements describing various boundary conditions (intuitively, the conditions that distinguish A-entities from B-entities of other sorts) and bridge principles if necessary. Bridge principles are empirically supported premises that connect the vocabularies of theories that do not share the same stock of predicates and terms. On Nagel's account, bridge principles are necessary for reduction if the reduced theory's vocabulary has predicates and terms that the reducing theory's vocabulary lacks. Suppose, for instance, that LA is a law statement of TA that is slated for derivation from LB, a law statement of TB:

LA For any x, if x is A_1, then x is A_2.
LB For any x, if x is B_1, then x is B_2.

Since the vocabulary of TB does not include the predicates A_1 or A_2, additional premises such as the following are required for the derivation:

ID_1 $A_1 = B_1$.
ID_2 $A_2 = B_2$.

Given ID_1 and ID_2, LA can be derived from LB by the substitution rule for identity.

The reduction of thermodynamics to statistical mechanics is often cited as an example of reduction via bridge principles. The term 'heat,' which occurs in the law statements of thermodynamics, is not included in the vocabulary of statistical mechanics. As a result, the derivation of thermodynamic law statements from mechanical ones requires the use of additional premises connecting the theories' respective vocabularies. An example might be the following:

Heat = mean molecular kinetic energy.

On a Nagelian account of reduction, theoretical identifications of this sort operate as bridge principles linking the vocabularies of reduced theories with the vocabularies of reducing theories.

One important criticism of Nagel's original account of reduction targeted his understanding of bridge principles. Nagel's original account did not require that bridge principles take the form of identity statements. Critics argued that this was a mistake (Sklar 1967, Schaffner 1967, Causey 1977, Hooker 1981). The idea that reduction involves the inclusion of one domain in another implies that the entities postulated by the reduced theory be identical to entities postulated by the reducing theory. In claiming to have reduced Kepler's laws to Newton's, the assumption is that planets are identical to massive bodies, not merely objects whose behaviors are correlated with the behaviors of massive bodies. If bridge principles are not identity statements like ID_1 and ID_2, then we do not have a case in which one domain is discovered to be part of another, more inclusive domain; we do not have a case in which the laws of the reduced theory can be explained by appeal to the laws of the

reducing theory, and so we do not have a case of explanatory consolidation, in which one theory takes over the descriptive and explanatory roles of another.

Lawrence Sklar (1967) argued that bridge principles must be identity statements by appeal to an empirical example. The Wiedemann-Franz law expresses a correlation between thermal conductivity and electrical conductivity in metals. It allows for the derivation of law statements about the latter from law statements about the former. This derivability, however, has never been understood to warrant the claim that the theory of electrical conductivity is reducible to the theory of heat conductivity, or vice versa. Rather, it points in the direction of a different reduction, the reduction of the macroscopic theory of matter to the microscopic theory of matter: the electrical conductivity of metals and the thermal conductivity of metals are both explainable by appeal to the properties of atoms and subatomic particles.

To illustrate in a general way the necessity of intertheoretic identities for reduction, imagine that Domains A and B comprise completely distinct entities whose behaviors are nevertheless correlated with each other. It turns out, for instance, that the principles governing A-entities and those governing B-entities mirror each other: for every A-law there is a corresponding B-law and vice versa, and instances of A-properties are correlated one-one with instances of B-properties. Because A-principles and A-properties mirror B-principles and B-properties, biconditional sentences such as the following end up being true:

BC_1 Necessarily, for any x, x is A_1 if and only if x is B_1;
BC_2 Necessarily, for any x, x is A_2 if and only if x is B_2.

Such biconditionals can underwrite a derivation of law statements such as LA from law statements such as LB: If BC_1 and BC_2 are both true, it is possible to derive LA from LB. What these biconditionals cannot underwrite, however, is the claim that TA is reducible to TB. The reason is that A and B are completely distinct domains, and an explanation of A-behavior must appeal to the principles governing A-behavior, not the principles governing B-behavior, even if the latter principles mirror the former. By analogy, if person *a* is killed by *c*, and person *b* is killed by *c**, the reason *a* died is that *c* killed him, not that *c** did, even if the circumstances in which *c** killed *b* mirror perfectly the circumstances in which *c* killed *a*. Explaining the behavior of A-entities requires that we get at the reasons why A-entities themselves behave as they do. If B-entities are entirely distinct from As, then appealing to the principles governing B-behavior doesn't accomplish this, not even if those principles are correlated with the principles governing As.

When applied to psychological discourse, the foregoing points imply that the reduction of psychological discourse to some branch of physical science requires that mental entities be identified with entities postulated by that branch of physical science. Psychophysical reduction could not involve two distinct yet coordinate domains. This is clear if we imagine a case involving psychophysical parallelism. Imagine two completely distinct ontological domains, one comprising bodies, the

other nonphysical Cartesian egos. Imagine, moreover, that these domains happen to mirror each other in the sense described above: the laws governing the behavior of bodies parallel the laws governing the behavior of the Cartesian egos, and the states of the Cartesian egos are distinct from, but nevertheless correlated one-one with, the states of bodies. In that case, it would be possible to derive statements about the behavior of Cartesian egos from statements about the behavior of bodies. This derivability, however, would not warrant the claim that the behavior of Cartesian egos was reducible to the behavior of bodies. The behavior of bodies could perhaps provide a helpful model or heuristic for understanding or predicting the behavior of Cartesian egos, but it would not provide a reducing theory which explained why the laws governing Cartesian egos obtained. Explaining the behavior of the egos requires that we get at the reasons why the egos behave as they do. If bodies are entirely distinct from the egos, then appealing to the principles governing bodies doesn't accomplish this, not even if those principles are correlated with the principles governing the egos. Cartesian egos act as they do on account of the laws governing their behavior, not on account of the laws governing the behavior of completely different entities. The same point would follow if some type of neutral monism were true—if, say, mental and physical phenomena were correlated with each other, but were both reducible to some third conceptual framework which was neither mental nor physical, but neutral. Mere correlations between instances of mental properties and instances of physical properties—even correlations that are law-like—are not sufficient to underwrite psychophysical reduction. Psychophysical reduction requires that mental properties be identical to physical properties.

There have been many other criticisms of Nagel's account of reduction. The syntactic model of theories, for instance, is widely regarded as false (Suppe 1974), as is the derivational model of explanation (Salmon 1989). It is nevertheless possible to formulate a more general account of reduction that captures the central insights of Nagel's account while avoiding its inadequacies. In general, we can say that theories are conceptual frameworks, systems of abstract representations, sentential or otherwise, that we use to describe and explain the behavior of entities in a given domain (Nersessian 1992; Trumpler 1997; Craver 2002). Explanation need not be understood to consist in derivation from law statements, and correspondingly, reduction needn't be understood to consist in the derivation of the reduced theory's law statements from the law statements of the reducing theory. The central feature of reduction is the takeover of descriptive and explanatory roles. This descriptive and explanatory takeover could take the form of derivations, but it could take other forms as well. What it requires is that the kinds of entities postulated by the reduced theory be identified with the kinds of entities postulated by the reducing theory. These theoretic identifications imply in turn systematic correlations between the representations of the reduced theory and representations of the reducing one, and these systematic correlations allow for descriptive and explanatory takeover. This generalized description of reduction should be sufficient for our purposes.

11.4 Nonreductive Physicalism

Now that we have a working understanding of reduction, we can consider nonreductive physicalism and whether the definition of physicalism proposed in Section 11.2 is compatible with it. Nonreductive physicalism accepts physicalism but rejects the reducibility of special sciences such as biology, psychology, and economics to physics; it denies, in other words, that physics can take over the descriptive and explanatory jobs the special sciences perform. How can this be if physicalism implies the core physicalist thesis? After all, the core physicalist thesis implies that all properties are physical, but if all properties are physical—if, that is, they are properties postulated by physics—then it seems difficult to avoid the conclusion that physics can perform all the descriptive and explanatory jobs that need performing.

We saw in Chapter 3 that Davidson (1974: 239) provides one way of understanding the nonreductive physicalist idea. According to Davidson, psychological discourse is irreducible to physical theory, not because there are two different kinds of entities, physical and other, but because we employ different kinds of conceptual schemes which carry with them different and 'disparate' commitments (Davidson 1970: 222). The vocabulary of a special science like psychology is crafted to satisfy descriptive and explanatory interests of one sort, and the vocabulary of physics is crafted to satisfy descriptive and explanatory interests of another sort. Because these vocabularies respond to different descriptive and explanatory interests, the vocabulary of the special sciences and the vocabulary of physics are not related in systematic ways. Fodor describes the idea as follows:

> Why ... should not the natural kind predicates of the special sciences cross-classify the physical natural kinds? Physics develops the taxonomy of its subject-matter which best suits its purposes: the formulation of exceptionless laws which are basic ... But this is not the only taxonomy ... [T]here are special sciences, with their specialized taxonomies ... [A]ll such taxonomies must apply *to the same things*. If physics is to be basic science, then each of these things had better be a physical thing. But it is not further required that the taxonomies which the special sciences employ must themselves reduce to the taxonomy of physics. (1974: 114)

The lack of systematic correlations between the vocabulary of a special science and the vocabulary of physics makes it impossible to identify the types of entities postulated by the special science with the types of entities postulated by physics. As a result, physics cannot take over the descriptive and explanatory roles the special science plays. Nonreductive physicalism is compatible with saying in this or that particular case that a given special scientific term refers to this entity or this collection of entities postulated by physics, or that this special scientific predicate is true of something on account of it having a property or standing in a relation postulated by physics. But this is not sufficient for reduction. Reduction involves systematic takeover of descriptive and explanatory roles. Systematic takeover of these roles

requires, however, that there be systematic correlations between the vocabularies of the special sciences and the vocabulary of physics, and nonreductive physicalism denies that there are such correlations.[4]

The foregoing points should make it clear why the proposed definition of physicalism is not committed to reductivism. Reduction involves descriptive and explanatory takeover. We saw in Section 11.3 that this requires that the entities postulated by the reduced theory be identified with the entities postulated by the reducing theory, and this requires in turn that there be systematic correlations between the vocabulary of the reduced theory and the vocabulary of the reducing one. The proposed definition of physicalism, however, does not imply that there are any such correlations. Even if everything is physical, the special sciences might employ vocabularies that are designed to satisfy descriptive and explanatory interests other than those that can be satisfied by physics. As a result, these vocabularies will not be systematically correlated with the vocabulary of physics, but will "cross-classify" that vocabulary, as Fodor puts it. If, however, the vocabularies of the special sciences cross-classify the vocabulary of physics, then physics will not be able to take over the descriptive and explanatory jobs the special sciences perform. The special sciences will thus be irreducible to physics. The proposed definition of physicalism is compatible, therefore, with the irreducibility of the special sciences; it is compatible, in other words, with nonreductive physicalism, and because of that, it does not fail to satisfy the second criterion for an adequate definition of physicalism discussed in Section 11.2, namely the criterion that a definition of physicalism not imply a commitment to a specific variety of physicalism such as reductivism.

The foregoing understanding of nonreductive physicalism stands opposed to an understanding that takes nonreductive physicalism to imply property dualism. We saw an example of the latter in Section 3.3, when we considered Cynthia Macdonald's remarks on nonreductive monism:

> Non-reductive monism is the view that each mental event is a physical event although mental properties are neither reducible to nor correlated in a . . . lawlike way to physical ones. [T]his theory seems . . . to reconcile monism at the level of particular events and their causal transactions, with the *sui generis* distinctness of the mental and physical at the level of properties. (1998: 346)

Another example is provided by Jaegwon Kim:

> *Property Dualism, or Nonreductive Physicalism*. The psychological properties of a system are distinct from, and irreducible to, its physical properties. (2006a: 13)

In Section 3.3, we noted one of the challenges that definitions like this face. Physicalism is a form of monism: it implies that everything is of only one kind. This idea is expressed by the core physicalist thesis: everything is physical. If we assume an

[4] For more on this point, see Jaworski 2011: 145–8.

ontology of individuals, properties, and events, the core physicalist thesis implies that all individuals, properties, and events are physical. This is incompatible with the claim that psychological properties are distinct from physical properties, for if psychological properties exist at all, physicalism implies that they must be identical to physical properties. If the psychological properties of a system really are distinct from its physical properties, as Kim's definition says, then the system's psychological properties must not be physical properties. But if they are not physical properties, then not everything is physical. Kim's definition of nonreductive physicalism is thus incompatible with the core physicalist thesis: what he calls 'nonreductive physicalism' cannot qualify by our lights as a form of physicalism at all; it is instead a form of DAT.

Macdonald's and Kim's definitions bring out another point as well. They attribute irreducibility to properties. Based on what we've said about reduction, however, this is a category mistake. Properties are not reducible or irreducible to other properties; rather, it is theories that are reducible or irreducible to other theories. To speak of the reducibility or irreducibility of A-properties to B-properties can be at best an elliptical way of speaking of the reducibility or irreducibility of the theory that postulates A-properties to the theory that postulates B-properties (Churchland 1986: 278). The relation between properties that is relevant to reduction is simply identity or non-identity. We saw in Section 11.3 that reduction requires that the entities—including properties—postulated by the reduced theory be identical to the entities postulated by the reducing theory. If A-properties are not identical to B-properties, then the theory that postulates A-properties is not reducible to the theory that postulates B-properties.

Definitions of nonreductive physicalism that appeal to property dualism are often motivated by two related ideas about properties. The first is a commitment to higher-order properties; the second, a commitment to abundant properties. Higher-order properties, recall, are logical constructions that quantify over other properties (Section 5.3). If F_1, $F_2, \dots,$ F_n are properties, we might define a property, H, by stipulating that something has H if and only if it has *some* F-property or other, or we might stipulate that something has H if and only if it has some F-property or other that satisfies some condition C. Exponents of higher-order properties sometimes understand nonreductive physicalism as the conjunction of physicalism with the claim that special scientific properties are higher-order properties (Sections 5.3 and 11.4). This understanding implies that there are two kinds of properties: the first-order properties postulated by physics and higher-order properties postulated by the special sciences. It thus suggests a kind of property dualism. This dualism, however, is not supposed to compromise a commitment to physicalism because special scientific properties are merely abstract postulates. To appreciate this point, suppose that P_1, $P_2, \dots,$ P_n are all the first-order physical properties that exist, and that we define a higher-order property, pain, as follows:

> Something has pain if and only if it has a physical property that satisfies the following condition: it results in the system wincing or groaning with a probability of $.N$ if the system receives pinpricks, burns, or abrasions.

By defining pain in this way, we do not add to the inventory of basic physical properties that exist. We cannot bring about the existence of new fundamental physical properties by sheer fiat—the way we can imagine God, say, bringing about the existence of a new fundamental physical property to add to $P_1, P_2, \ldots, P_n$. Defining a new higher-order property does not add to the basic physical properties of the world, it merely introduces a new way of talking about those properties. Having pain amounts to having some physical property or other that satisfies the aforementioned condition. Saying that something has pain, therefore, is just another way of describing its possession of physical properties—a way that abstracts from the kinds of distinctions made by the vocabulary of physics.

Critics have nevertheless argued that this implication of nonreductive physicalism is a liability. In Section 5.3, we saw that both Kim (1998) and Prior et al. (1982) argue that higher-order properties can do no causal work other than the work done by their lower-order realizers (their *causal bases* in Prior et al.'s terminology). If properties are sparse and the Eleatic principle discussed in Section 2.2 is true, then so-called second-order properties are not real properties; they are simply predicates that apply to things on account of those things having certain first-order properties. If pain is a so-called second-order property, and second-order properties are real properties—that is, sparse properties with causal essences—then pain would have to have a causal essence. That essence, moreover, would have to be distinct from the essence of every other property in order for pain to be distinct from every other property. But if pain is second order, then the causal powers of this or that particular instance of pain couldn't be any different from the causal powers of its lower-order realizer. If, for instance, an individual *a* were in pain at *t* on account of having property P_1 at *t*, the causal powers of *a*'s pain could not be any different from the causal powers of P_1. From what source, after all, could *a*'s pain acquire causal powers in addition to those of P_1 if pain is a so-called second-order property? Surely the mere application of a predicate cannot bring new causal powers into existence. But in that case, pain would either have to be identical to P_1, or else we would have to conclude that it was not a real property at all. If pain is multiply realizable, moreover, as nonreductive physicalists typically maintain, then it cannot be identical to P_1, for if pain is multiply realizable, it is possible for something to have pain without having P_1. Something could have pain by virtue of having some other property satisfying condition C. Pain, therefore, must not be a real property if it is second order. And the same is true of other so-called second-order properties. It is for precisely this reason that Kim (1998: 104) has claimed that it is misleading to speak of higher-order properties as opposed to higher-order predicates, concepts, or descriptions. From this perspective, the nonreductive physicalist view described above is committed to there being higher-order predicates, but not to there being higher-order properties, for there are no such properties. Correspondingly, from this perspective, nonreductive physicalism should not be described as a form of property dualism. The term 'property dualism' should instead be reserved for a dualism of first-order properties.

I've argued in this section that nonreductive physicalism should be understood in the way Davidson and Fodor suggest: everything is physical; everything can be exhaustively described and explained by physics, yet the special sciences employ vocabularies that satisfy descriptive and explanatory interests other than those that can be satisfied by physics. It should be clear that this understanding is committed to two different conceptions of description and explanation. One is operative in the claim that everything can be described and explained exhaustively by physics, the other in the claim that physics cannot perform the descriptive and explanatory jobs the special sciences do. The first conception corresponds to what we might call the objective factors that contribute to this or that effect. If physicalism is true, physics is capable in principle of describing all such factors. The second conception corresponds to the subjective interests that underwrite intuitive judgments about what kinds of descriptions and explanations are satisfactory. One implication of nonreductive physicalism is that descriptions and explanations of the latter sort needn't correspond to descriptions and explanations of the former sort, that the kinds of descriptions and explanations we find intuitively satisfying needn't correspond to the objective factors that are responsible for this or that effect. In fact, if nonreductive physicalism is true, it is difficult to see how the descriptions and explanations of the special sciences could correspond to the objective factors that are responsible for this or that effect. For according to nonreductive physicalism, the descriptions and explanations physics provides do not satisfy the interests that motivate the special sciences. Satisfying those interests requires vocabularies that track factors different from those tracked by the vocabulary of physics. But if physicalism is true, the objective factors that are responsible for producing particular effects are precisely the ones tracked by the vocabulary of physics.

The foregoing account of explanation marks one important difference between nonreductive physicalism and hylomorphism. In Chapters 8 and 10, I outlined an account of explanation and its relation to objective causal factors that takes a cue from John Stuart Mill's concept of a complete cause (1843: Book 3, chapter 5, section 3: 398ff.). On that account, any given event owes its occurrence to a wide range of different factors, which together constitute the event's complete cause. In any given context we pick and choose which factors belonging to the complete cause interest us. As David Lewis says:

> We sometimes single out one among all the causes of some event and call it "the" cause, as if there were no others. Or we single out a few as the "causes," calling the rest mere "causal factors" or "causal conditions." Or we speak of the "decisive" or "real" or "principal" cause. We select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we want to talk about. (1973b: 162)[5]

[5] Elsewhere, Lewis expresses the idea as follows: "The multiplicity of causes and the complexity of causal histories are obscured when we speak, as we sometimes do, of *the* cause of something. That suggests that there is only one . . . If someone says that the bald tyre was the cause of the crash, another says that the

It typically takes a whole range of different factors to cause a car crash: the balding tires, the faulty brake mechanism, the inadequate roadway grading, the driver's blood-alcohol level—all of these and more contribute to the crash, all of them belong to its complete cause. In one context or another, however, we focus on some of these contributing factors and ignore the rest. The automotive engineer focuses on the brake mechanism, the civil engineer on the roadway grading, the prosecutor on the blood-alcohol level, and so on. These foci are expressed in different explanations of the event. The automotive engineer says that the crash happened because of the brake mechanism, the civil engineer says it happened because the roadway wasn't graded steeply enough, etc. These different acts of explaining mark a selection from among the wide range of causes that operate in a given case. In any given case, moreover, we can trace the basis of the selection that's been made. When we ask why the crash happened, the automotive and the civil engineer provide different answers because they are answering different questions. The expression, "Why did this car crash?" has a complex logic that is hidden beneath its simple surface grammar—a logic that Belnap and Steel (1976) and van Fraassen (1980) have done much to reveal. The automotive engineer looks to answer a question like Q1; the civil engineer, a question like Q2:

Q1. Why did this car crash when cars with different brake mechanisms didn't crash under similar conditions?
Q2. Why did this car crash when cars on roadways with different grades didn't crash under similar conditions?

These questions have different contrast classes; they request different kinds of information. In each case, that information is to be found among the objective factors that contributed to the crash. Q1 and Q2 are not requests to fabricate an account of the crash; they are requests to discover which factors, among those that actually contributed to the crash, supply the relevant information.

There is nothing subjective about explanation on this account. Our acts of explanation pick out objective features of the world. What is up to us, or subjective, are the motivations that lead us to make a selection. What our selections are selections of, however, are objective contributing factors. If the world consists of individuals having particularized properties, then the explanation that the car crashed at *t* because its brakes were faulty picks out mind-independent constituents of the world and a certain kind of causal relation between them. As a result, there is

driver's drunkenness was the cause, and still another says that the cause was the bad upbringing that made him reckless, I do not think that any of them disagree with me when I say that the causal history includes all three. They disagree only about which part of the causal history is most salient for the purposes of some particular enquiry . . . Some parts will be salient in some contexts, others in others. Some will not be salient in any likely context, but they belong to the causal history all the same: the availability of petrol, the birth of the driver's paternal grandmother, the building of the fatal road, the position and velocity of the car a split second before the impact" (Lewis 1986c: 215–16).

no disconnect on this account between the interests that motivate the descriptive and explanatory work of the special sciences and the objective factors that contribute to this or that effect. The interests that motivate the special sciences look to discover factors that contribute to what happens in the world in ways that differ in kind from the factors that contribute in the ways physics describes. According to hylomorphism, these differences in kind are grounded in structure. *Structure makes a difference*: it contributes to things in ways that differ fundamentally from the way unstructured things do. This difference between the hylomorphic account of explanation and the nonreductive physicalist one has implications for their respective approaches to the problem of downward causation, a topic I discuss in detail in Section 13.3.

11.5 Physicalism and Supervenience

Critics might object that the definition of physicalism proposed in Section 11.2 makes no mention of supervenience. Any adequate definition of physicalism must mention supervenience, they might say, since supervenience can plausibly be taken as a minimal physicalist commitment (Kim 2006a: 13). I will argue that this objection is misguided for two reasons. First, saying that supervenience is a minimal physicalist commitment means only that an adequate definition of physicalism must imply a commitment to it, it does not mean that a definition of physicalism must be formulated in terms of it. Second, definitions of physicalism that are formulated in terms of supervenience have difficulty accommodating the core physicalist thesis, whereas the proposed definition has no such difficulty. Since that definition also implies a commitment to supervenience, it is arguably superior to supervenience-based alternatives.

Recall that supervenience is a kind of property covariation. To say that A-properties supervene on B-properties is to say that things cannot differ in their A-properties without also differing in their B-properties, or equivalently, that B-twins must be A-twins.

Psychophysical supervenience is the thesis that mental properties supervene on physical properties. There are stronger and weaker versions of this claim:

Weak psychophysical supervenience: For any possible world w, and any x and y in w, if x and y are physically indistinguishable, then they are mentally indistinguishable.

Strong psychophysical supervenience: For any possible worlds w_1 and w_2, and any x in world w_1 and y in world w_2, if x and y are physically indistinguishable, then they are mentally indistinguishable.

The difference between these definitions is that weak psychophysical supervenience specifies a psychophysical covariation relation only within a world. In any one world, it says, physical indiscernibility guarantees mental indiscernibility. But knowing that mental and physical properties covary in one world tells us nothing about whether and/or how they covary in others. As a result, weak psychophysical supervenience is

compatible with the possibility that a physical replica of me might have radically different mental properties or perhaps no mental properties at all. Strong psychophysical supervenience rules this out. Because *x* and *y* can belong to different possible worlds, it specifies a covariation relation that obtains across worlds, not just within a single world. It says that for any individuals in any worlds at all—individual *x* in world_1 and individual *y* in world_2—physical indiscernibility guarantees mental indiscernibility. It is impossible for physical twins not to be mental twins, even if they are in different worlds. Strong psychophysical supervenience implies weak psychophysical supervenience, but not vice versa.

The definition of physicalism proposed in Section 11.2 implies strong psychophysical supervenience. To see this, suppose that physicalism is true. We saw in Section 11.2 that physicalism implies the core physicalist thesis; it thus implies that all properties are physical. Suppose, then, that *x* and *y* are physically indistinguishable; that is, *x* and *y* have exactly resembling physical properties. In that case, *x* and *y* will be indistinguishable tout court, for if physicalism is true and all properties are physical, then there will be no other properties in terms of which *x* and *y* can differ from each other. Consider mental properties in particular. Either they exist or they do not. If mental properties exist, and physicalism is true, then they must be identical to physical properties, for if physicalism is true, physical properties are the only properties there are. But if mental properties are identical to physical properties, then strong psychophysical supervenience follows straightaway. It will be impossible for *x* and *y* to differ from each other mentally without differing from each other physically, for if mental properties are identical to physical properties, a difference in mental properties can only be a difference in physical properties. Suppose, on the other hand, that physicalism is true and that mental properties do not exist. In that case, strong psychophysical supervenience follows trivially, for if mental properties do not exist, then necessarily two things cannot differ from each other mentally: there will be no mental properties in terms of which they can differ. Consequently, everything will be mentally indistinguishable from everything else: each thing will have exactly the same mental properties as every other thing, namely none. The consequent of strong psychophysical supervenience will thus be satisfied, so the entire thesis will be true.

The foregoing considerations show that the proposed definition of physicalism implies a commitment to strong psychophysical supervenience. Moreover, what is true of mental properties is true of other properties as well. Similar proofs can be constructed for properties of any sort one likes. Consequently, the proposed definition of physicalism implies the following general supervenience thesis:

> *Strong supervenience*: Necessarily, for any *x* in world w_1 and any *y* in world w_2, if *x* and *y* are physically indistinguishable, then they are indistinguishable simpliciter.

Given reasonable assumptions, moreover, the proposed definition also implies psychophysical necessitation:

> *Psychophysical necessitation*: Necessarily, for any *x*, if *x* has a mental property, *M*, then there is a physical property, *P*, such that *x* has *P*, and necessarily, for any *z*, if *z* has *P*, then *z* has *M*.

If physicalism is true, then either mental properties exist or they do not. If mental properties do not exist, then the antecedent of psychophysical necessitation goes unsatisfied, and the entire thesis is trivially true. If, on the other hand, mental properties do exist, then both the antecedent and the consequent of psychophysical necessitation are satisfied, and the entire thesis is true. If physicalism is true and mental properties exist, then those properties must be identical to physical properties, for if physicalism is true, all properties are physical. A property cannot exist without itself, so if *x* has a mental property, *M*, then it must have a physical property, *P*, namely the physical property to which *M* is identical. But if *M* is identical to *P*, then necessarily anything that has *P* will also have *M*: property identity implies necessary coextension. Consequently, psychophysical necessitation is implied by physicalism, whether mental properties exist or not.

Critics might see in the foregoing results the basis of a further objection to the definition of physicalism proposed in Section 11.2. Strong supervenience and/or necessitation have sometimes been taken to imply reductivism (Kim 1984: 73–4; 1989: 278–9; 1990: 150–1). Given certain assumptions about properties, if mental properties are either necessitated by or strongly supervenient upon physical properties, then mental properties and physical properties end up corresponding one-one. If one-one psychophysical correlations are sufficient for psychophysical reduction, then a commitment to strong psychophysical supervenience will imply a commitment to psychophysical reductivism, and if that is the case, then someone might object that the proposed definition fails to satisfy the second criterion of adequacy—the criterion that a definition of physicalism cannot imply a commitment to a specific variety of physicalism such as reductivism.

There are two things to say about this objection. First, it is based on a false assumption (McLaughlin 1995: 45–8). It assumes that one-one psychophysical correlations are sufficient for reduction, but we saw in Section 11.3 that this is not the case. Reduction requires the identification of entities postulated by the reduced theory with entities postulated by the reducing theory. One-one correlations by themselves are not enough to secure this. We saw in Section 11.3 that psychophysical parallelism illustrates that it is possible for entities in domain A to be correlated one-one with entities in domain B without the theory describing and explaining the behavior of A-entities being reducible to the theory describing and explaining the behavior of B-entities.

Second, Kim's argument that necessitation and/or strong supervenience imply one-one psychophysical correlations is based on the assumption that Boolean complementation and either infinitary conjunction or infinitary disjunction are property-forming operations. We saw in Section 9.1, however, that this assumption is dubious

if properties are sparse, for in that case, properties are things that confer powers on individuals, and it is far from evident that every predicate constructed through the application of Boolean operations must express a distinctive power-conferring property. For these reasons, psychophysical necessitation and strong psychophysical supervenience cannot be taken to imply psychophysical reduction.

11.6 Necessitation Physicalism and Abundant Properties

The definition of physicalism advanced in Section 11.2 presupposes that properties are sparse. We argued in Chapter 2 that there are good reasons to think that properties are sparse. There is nevertheless another approach to defining physicalism made popular by David Lewis (1983b), among others. It presupposes that properties are abundant. Unlike definitions of physicalism that presuppose sparse properties, definitions that presuppose abundant properties don't make notions of supervenience or necessitation mere implications of physicalism; they make those notions central to the very definition of physicalism. David Chalmers' definition provides an example of this approach:

> *Materialism* (or *physicalism*) . . . is true [in a world] if all the positive facts about the world are globally logically supervenient on the physical facts. (1996: 41)

Recall from Section 10.5 that according to Chalmers, A-properties globally logically supervene on B-properties exactly if, for any logically possible worlds w_1 and w_2 that are B-indiscernible from each other, the A-facts that are true of w_1 are also true of w_2.[6] Facts, moreover, are instantiations of properties. According to Chalmers, then, if physicalism is true in world w_1, then for any world w_2 whose instantiations of positive physical properties are indiscernible from the instantiations of positive physical properties in w_1, the instantiations of all positive properties in w_2 are indiscernible from the instantiations of all positive properties in w_1. Consider likewise Daniel Stoljar, whose definition appeals to necessitation instead of supervenience:

> [E]very instantiated property is necessitated by, and not metaphysically distinct from, some physical property. (2010: 235)

A property F necessitates a property G, according to Stoljar, exactly if necessarily, for any x, if x has F, then x has G. Physicalism thus implies that for any instantiated property, G, there is a physical property, F, such that necessarily anything that has F also has G.

[6] Chalmers (1996: 40) frames his own definition of global logical supervenience in terms of indiscernibility with the actual world. The definition I've offered represents a generalization of that definition to all worlds.

The reason notions of supervenience or necessitation are central to definitions of physicalism that presuppose abundant properties is that exponents of abundant properties are committed to more or less all predicates expressing properties. Since we in fact use biological, psychological, economic, and other special scientific predicates, they are committed to claiming that there are biological, psychological, economic, and other special scientific properties. According to them, what distinguishes these properties from physical properties is that the latter are in some way more fundamental. In order to define what fundamentality consists in, they appeal to notions of supervenience or necessitation: the fundamental properties are the ones on which all properties supervene or the properties by which all properties are necessitated. If, for instance, individual_1 and individual_2 have the same fundamental properties, then they must have the same properties tout court. Likewise, if possible world_1 and possible world_2 have the same distribution of fundamental properties over individuals, then they must have the same distribution of all properties over individuals. On these views, physicalism is the claim that the fundamental properties are all physical, and that everything is necessitated by them or by their distribution over individuals in a world.

Definitions of physicalism that presuppose sparse properties do not need a notion of fundamentality like this. If physicalism is true and properties are sparse, then the only properties that exist are physical. Physical properties are not merely a proper subset of all the properties, they are all the properties. Consequently, it is unnecessary to posit supervenience or necessitation relations that explain what makes the physical properties more fundamental than other properties, for if properties are sparse, physicalism implies that there are no properties other than physical ones. If biological, psychological, and other special scientific predicates express any properties at all, they must express physical properties. It is simply a matter of determining whether or not they do, and if so, how the semantics of those predicates manages to do the job.

Stoljar (2010: 40) has attacked sparse properties on two grounds. First, he says, only an abundant property theory allows for the kind of generality that is needed for an adequate definition of physicalism. This objection is nevertheless difficult to appreciate, for we have seen that if properties are sparse, physicalism can be defined as the claim that everything is physical, and it is difficult to see how this definition lacks generality; if anything, it seems much more general than Stoljar's own.

Second, Stoljar argues that a view committed to sparse properties would have to be committed to there being fundamental properties in Lewis' sense:

> This world, or any possible world consists of things which instantiate fundamental properties and which, in pairs or triples or . . . , instantiate fundamental relations. Few properties are fundamental: the property of being a club or a tub or a pub, for instance, is an unnatural gerrymander, a condition satisfied by miscellaneous things in miscellaneous ways. A fundamental, or 'perfectly natural,' property is the extreme opposite. Its instances share exactly some aspect of their intrinsic nature. Likewise for relations. I hold, as an a priori

principle, that every contingent truth must be made true, somehow, by the pattern of coinstantiation of fundamental properties and relations. The whole truth about the world, including the mental part of the world, supervenes on this pattern. If two possible worlds were exactly isomorphic in their patterns of coinstantiation of fundamental properties and relations, they would be exactly alike simpliciter. (Lewis 1994: 412)

According to Stoljar, Lewis' notion of fundamental properties is too speculative to ground a workable notion of sparse properties (2010: 35–6). It is empirically possible that there might be no fundamental properties, and it is metaphysically contentious whether the whole truth about a world should depend on the coinstantiation of fundamental properties. Consequently, Stoljar concludes that we should take properties to be abundant.

There are at least two problems with Stoljar's second argument. First, it appears to attack a straw man since a commitment to sparse properties does not entail a commitment to fundamental properties in Lewis' sense. Lewis introduces a notion of fundamentality only because he presupposes that properties are abundant, and thus needs a way of explaining what makes physical properties special if physicalism is true. What makes them special, according to Lewis, is that they are fundamental. We've seen, however, that if properties are sparse, introducing a notion of fundamentality becomes unnecessary. If properties are sparse and physicalism is true, we need not claim that physical properties are more fundamental than other properties, for if properties are sparse and physicalism is true, there are no other properties. By assimilating sparse properties to Lewis' fundamental properties, Stoljar overlooks a host of alternative accounts of sparse properties, including the kind of account developed in Chapters 2–5.

Second, if we take seriously Stoljar's premise that the speculative character of a thesis counts as a strike against it, then there is reason to favor sparse properties over abundant ones. Suppose that a theory's degree of speculativeness can be understood in terms of its commitments in a particular disputed domain: the more commitments a theory has in a disputed domain, the more speculative the theory is. On this measure of speculativeness, a sparse property theory seems less speculative than the abundant alternative. We discussed the reasons for this in Chapter 2: defending abundant properties against the objections requires a commitment to something like Lewis' extreme modal realism. In addition, we'll see momentarily that exponents of abundant properties must take on yet further commitments in order to defend a workable definition of physicalism.

The foregoing considerations show at the very least that Stoljar's objection fails to rule out definitions of physicalism that presuppose sparse properties. In addition, if we grant Stoljar's premise that a metaphysical framework's degree of speculativeness can count as a strike against it, then there is further reason to endorse sparse properties over abundant ones since a sparse property theory makes fewer speculations about, say, the metaphysics of modality than its abundant competitors (Sections 2.2–2.3).

In addition, definitions of physicalism that presuppose abundant properties have difficulty accommodating the core physicalist thesis without appealing to contentious metaphysical assumptions. The reason is that it seems possible that nonphysical properties might supervene on or be necessitated by physical properties. Imagine an emergentist DAT similar to C. D. Broad's (1925). It claims that there are physical properties and distinct nonphysical mental ones, and that the two kinds of properties are connected by psychophysical laws. Imagine that such laws are metaphysically necessary; whenever something has a nonphysical mental property, *M*, it has a specific physical property, *P*, and anything that has *P* has *M* in all metaphysically possible worlds. Suppose, moreover, that there is no way to bring about an instance of *M* other than by bringing about an instance of *P*. According to such a theory, there could, as a matter of metaphysical necessity, be no mental differences without corresponding physical differences. A certain distribution of physical properties would necessitate a certain distribution of mental ones. Yet, not everything would be physical since there would be nonphysical properties.

Stoljar calls this kind of view 'necessitation dualism.' To distinguish physicalism from necessitation dualism, defenders of abundant properties must either introduce conditions other than mere supervenience or necessitation to define physicalism, or else they must rule out necessitation dualism on the grounds that it is incoherent. Stoljar discusses one strategy of the latter sort. It involves stipulating that there can be no metaphysically necessary psychophysical laws, a stipulation that takes the form of Hume's dictum, the claim that there can be no necessary connections between metaphysically distinct existences—in this case, between metaphysically distinct properties. If Hume's dictum is true, then it is impossible for there to be metaphysically necessary connections between physical properties and distinct mental ones, as necessitation dualism claims. Necessitation dualism is thus an incoherent view, and so poses no challenge to definitions like Stoljar's.

The problem with this argument is that there appear to be counterexamples to Hume's dictum. Mumford (2004: 166–9) mentions at least three different kinds.[7] First, there appear to be necessary connections between being colored and having a shape:

> Colour requires an area to be coloured, but every area is bounded. A bounded area must be of some shape. There is thus a chain of necessary connections through colour, area, boundary, to shape. At no point does the chain appear weak. And we cannot reasonably maintain that the necessary connections are merely logical or analytic. They must be grounded in the natures of the properties involved... Something may change its colour while retaining its shape, or vice versa. These two properties thus qualify as distinct existences. (Mumford 2004: 167)

[7] Mumford nevertheless appears to have moved away from the idea that there is a necessary connection between a power and its manifestation. See, for instance, Mumford and Anjum 2011.

Second, there are necessary exclusions among properties. Mumford's example is the relation among the determinates of a determinable. If something is colored, then it must be a determinate hue such as red, blue, or yellow. If something is one hue, however, then it cannot also be another hue. Necessarily, something that is red all over cannot also be blue all over. Being red and being blue appear to be distinct existences; consequently, it looks as though there can be necessary connections between distinct existences. Finally, Mumford gives the example of powers and their manifestations. These appear to be distinct existences since it is possible for something to have a power without manifesting it. Yet, there is a necessary connection between a power and its manifestation: a fragile object, for instance, must break if struck the right way; it is metaphysically impossible for something to be fragile and not breakable.

These examples give us reason to think that Hume's dictum is false.[8] Are there any reasons to think that it is true? Stoljar cites David Lewis and David Armstrong as defenders of it. Lewis (1986d: 101–2) defends it as part of his assault on structural universals. Denying Hume's dictum, he suggests, is tantamount to believing in magic:

> Why *must* it be that if something instantiates *methane*, then part of it must instantiate *carbon*? . . . [O]n the present conception, this necessary connection is just a brute modal fact- . . . Therein lies the magic . . . If you said that whenever *carbon* is instantiated, *bromine* must necessarily be instantiated next to it, that would make good enough sense as a matter of nomological necessity. There is no such law of nature, but there could have been. But suppose you said that it was a matter of necessity *simpliciter*—absolute 'logical' or 'metaphysical' necessity. Then what you say is not only false; it is entirely unintelligible how it could be true. Why couldn't anything over here coexist with anything else over there[?] . . . [T]he universal *carbon* has nothing more in common with the universal *methane* than *rubidium* has! They are three distinct atomic individuals, and that is that. There is no conceivable reason why the universal *methane* should, by the strictest necessity, drag the universal *rubidium* around with it wherever it goes. How does it manage, then, to drag around *carbon*? . . . Although we understand just what necessary connections are supposed to obtain, we are given no notion how they possibly could. (1986d: 101–2)

Lewis claims that the idea of necessary connections between distinct existences is unintelligible. He nevertheless does not address the kinds of examples Mumford cites, ones which do not appear to pose any challenge to intelligibility. In Chapter 5, we saw that Lewis is a categoricalist: he denies that properties have dispositional essences. Consequently, it should not surprise us if he rules out the example of powers. But denying the existence of powers falls short of establishing that their existence is unintelligible, nor does Lewis address the other kinds of examples Mumford cites. His argument thus appears to fall short of providing a convincing argument in favor of Hume's dictum.

[8] See Wilson 2010 for a more detailed discussion and criticism of Hume's dictum and the arguments for it.

Armstrong (1989b: 116–18) advances a different argument. Hume's dictum, he says, is entailed by the combinatorial theory of possibility he espouses. Since powers are ostensibly counterexamples to Hume's dictum, Armstrong concludes that powers must not exist: all properties must instead be categorical. Although this is a more substantial line of reasoning than Lewis', it depends on a prior commitment to Armstrong's combinatorialism. If Armstrong argues by *modus ponens* from his combinatorialism to the conclusion that there are no powers, defenders of powers can argue by *modus tollens* that since there are powers, Armstrong's combinatorialism must be false. Armstrong's argument is thus inconclusive at best, and in later work he seems to acknowledge as much: "There seems to be no conclusive argument against such a position," he says (1997: 99).

The foregoing considerations show that the attempt to rule out necessitation dualism relies on a highly contentious metaphysical thesis, one for which there appear to be counterexamples. Stoljar suggests, however, that these considerations might attack a straw man. They assume that 'metaphysically distinct' means 'numerically distinct,' that F is metaphysically distinct from G exactly if F is not identical to G. But this cannot be what necessitation dualists mean by 'metaphysically distinct,' Stoljar insists; they must instead be committed to mental and physical properties being distinct in a sense more robust than mere numerical distinctness. The reason, he says, is that the mere numerical distinctness of mental and physical properties is compatible with physicalism, in particular, with a nonreductive physicalist view that endorses multiple realizability. On this kind of view, it is possible for instances of a given mental property, M, to be correlated with instances of different kinds of physical properties, $P_1, P_2, \ldots, P_n$. If M were identical to one of these physical properties, P_i, then necessarily every instance of M would be correlated with an instance of P_i, so if it were possible for something to have M without having P_i, as the multiple realizability thesis maintains, then M would not be identical to P_i, and the same would be true mutatis mutandis of other physical properties. Yet, Stoljar maintains, this nonidentity of mental and physical properties is compatible with physicalism. Consequently, the numerical distinctness of mental and physical properties is insufficient to capture the kind of metaphysical distinctness that necessitation dualists endorse: they must be committed to a notion of metaphysical distinctness more robust than mere numerical distinctness.

The problem with Stoljar's argument is that it hinges on the assumption that the multiple realizability thesis implies a dualism of properties that is compatible with physicalism. The considerations advanced in Section 11.4, however, provide good reasons for thinking that this assumption is false. If the multiple realizability thesis implies property dualism, then it must imply either a dualism of first-order properties alone or else a dualism that posits higher-order properties of some sort. But Stoljar cannot take the multiple realizability thesis to imply a dualism of first-order properties alone since that kind of property dualism is incompatible with physicalism. If the multiple realizability thesis implies a dualism of first-order properties, then

there is a first-order property *M* that is not identical to any first-order physical property; *M*, in other words, is a nonphysical property. But the existence of a nonphysical property is incompatible with the core physicalist thesis. Consequently, if the multiple realizability thesis implies property dualism, then it must be a dualism that posits higher-order properties—a dualism, for instance, according to which mental properties are higher-order properties distinct from the lower-order physical properties over which their definitions quantify. The problem with this latter proposal, as we saw in Section 11.4, is that there is good reason to think that there are no higher-order properties, and if there are no such properties, then the multiple realizability thesis cannot be taken to imply property dualism of this higher-order sort either. As a result, the multiple realizability thesis cannot be taken to support a dualism of properties, at least not a dualism of properties that is compatible with physicalism. Stoljar's argument thus trades on a false assumption. It gives us no reason to think that necessitation dualists must take metaphysical distinctness to be anything but numerical distinctness—the numerical distinctness of first-order properties in particular.

Someone might object at this point that if I am right, there is no coherent way of understanding a nonreductive physicalist view that endorses multiple realizability, and that this constitutes a reductio ad absurdum of my position. But this objection is based on a false premise. We saw in Section 11.4 that there is a coherent way of understanding nonreductive physicalist views at large, including views that endorse multiple realizability. According to that understanding, a nonreductive physicalist view would claim that mental predicates (not properties) are higher-order predicates, and that because of this, their deployment is not correlated in a systematic way with the deployment of physical predicates. This is an understanding of nonreductive physicalism in terms of a dualism of mental and physical predicates or concepts, not a dualism of mental and physical properties. Reductive physicalist theories are also committed to a dualism of mental and physical predicates or concepts since they claim that mental predicates and physical predicates express the same properties. They nevertheless differ from nonreductivist theories because they claim that mental and physical predicates or concepts are correlated in systematic ways, ones that allow for the identification of mental properties with physical properties. Nonreductivists deny that there are systematic correlations of this sort, and the multiple realizability thesis is one form this denial can take.

Based on the foregoing considerations, I conclude that the attempt to rule out necessitation dualism fails. Necessitation dualism poses a challenge to definitions of physicalism that appeal to supervenience or necessitation alone. Stoljar's own definition, however, does not. He adopts the first strategy mentioned above. He adds a condition to his definition of physicalism: properties are not only necessitated by physical properties, his definition says, they are also not metaphysically distinct from physical properties. Stoljar discusses four different ways of defining 'metaphysically distinct,' so his definition can be understood in at least four different ways:

(1) Every instantiated property is necessitated by, and is not numerically distinct from, some physical property.
(2) Every instantiated property is necessitated by, and is not distinct in essence from, some physical property.
(3) Every instantiated property is necessitated by, and is not modally distinct from, some physical property.
(4) Every instantiated property is necessitated by, and is not primitively distinct from, some physical property.

I will argue that each of these ways of understanding Stoljar's definition is problematic.

The problems with (1) are twofold. First, the numerical distinctness condition appears to render the necessitation condition otiose. If no instantiated properties are numerically distinct from physical properties, then all properties are numerically identical to physical properties. In that case, though, the necessitation condition follows trivially. If *F* is identical to *G*, then ipso facto anything that has *F* must have *G*. If all properties are identical to physical properties, all properties will be necessitated by physical properties. It thus becomes unnecessary to add necessitation as a further condition. Second, (1) implies that all properties are physical, but this implication is problematic for at least two reasons. First, if all properties are physical, then Stoljar's definition seems little different from the definition proposed in Section 11.2. And since the latter definition is less metaphysically contentious, there is good reason to favor it over Stoljar's. Second, Stoljar takes the claim that all properties are physical to be implausibly strong (2010: 33–4). Consider, for instance, the properties of the U.S. Supreme Court, such as having the power to overturn laws passed by Congress. Surely having this power is a legal or social property, not a physical one, says Stoljar. It is, after all, not a property postulated by physics, and in that case, it seems implausible to consider it a physical property. Yet, as implausible as Stoljar takes this result to be, it is what he is committed to if he endorses (1). Notice that exponents of sparse properties do not face this difficulty. They would say that the predicate 'has the power to overturn laws passed by Congress' is a legal or social predicate. This claim by itself has no implications for whether having the power to overturn laws passed by Congress is a property. Nor does it have implications for what kind of property it might be. The predicate 'has the power to overturn laws passed by Congress' might express a (first-order) legal or social property distinct from any physical property (a view that rejects physicalism), or it might express a physical property—the same physical property each time it is applied (a reductive physicalist view), or different physical properties on different occasions (a nonreductive physicalist view)—or it might express no property at all (an eliminativist view). But the division between predicates and properties that affords exponents of sparse properties this flexibility is not available to exponents of abundant properties like Stoljar. He is committed to 'has the power to overturn laws passed by Congress' expressing a property, and according to (1), that property must be physical.

Consider now (2). It claims that, according to physicalism, every instantiated property is necessitated by some physical property, and has the same essence as that physical property. The problem with this definition is that it appears either to collapse into (1) or else to be unintelligible. To see this, let us assume first that if property *F* and property *G* have the same essence, then *F* is numerically identical to *G*. In that case, (2) implies that all properties are numerically identical to physical properties, just as (1) says, for according to (2), every instantiated property has the same essence as some physical property, and by our assumption, if property *F* and property *G* have the same essence, then they must be one and the same property. Suppose, on the other hand, that our assumption is false, that it is possible for *F* and *G* to have the same essence, and yet to be numerically different properties. This claim is difficult to understand. It is easy to understand how there could be two different individuals that share the same essence. Betty and Sally might both be essentially human, and yet still differ from each other on account of having different nonessential properties. But it is difficult to imagine an analogue for the case of properties. What, after all, could distinguish one property from another if not their essences? If property *F* and property *G* have the same essence, then something else must distinguish *F* from *G*, and whatever that something else is, it must distinguish them from each other in all possible worlds in which they exist. But how could something distinguish *F* from *G* in all possible worlds in which they exist and yet not be essential to them? In the absence of clear answers to these questions, (2) remains speculative at best and unintelligible at worst.

Consider now (3). To say that *F* and *G* are modally distinct is to say that *F* can be instantiated without *G* and vice versa. So according to (3), physicalism is the claim that every instantiated property is necessitated by some physical property and cannot be instantiated apart from that physical property. The problem with (3) is that it is compatible with necessitation dualism. Necessitation dualists are free to claim that no properties can be instantiated without physical properties being instantiated. Consequently, if Stoljar's definition of physicalism is interpreted as (3), then it fails to imply the core physicalist thesis: it fails to imply that everything is physical.

Consider finally (4). It takes metaphysical distinctness to be a primitive notion which is not further explicable by appeal to other notions. The problem with (4) is that it appears either to collapse into (1) or else not to imply the core physicalist thesis. (4) implies that there are no properties which are primitively distinct from physical properties. What exactly does (4)'s denial of primitive distinctness imply? In particular, does it imply the numerical distinctness of properties? If *F* is primitively distinct from *G*, does that imply that *F* is numerically distinct from *G* or not? If it does, then (4) collapses into (1), for in that case, (4) implies that all properties are numerically identical to physical properties. If, on the other hand, (4) does not imply this, then it is compatible with there being properties that are not identical to physical properties, but this latter claim is incompatible with the core physicalist thesis. It

seems, then, that however Stoljar's definition of physicalism is understood, the result is problematic.

What the foregoing arguments appear to show is that exponents of abundant properties, like Stoljar, will have difficulty defining physicalism in a way that both distinguishes it from necessitation dualism and that also avoids making physicalism implausibly strong. The definition of physicalism proposed in Section 11.2 faces no analogous difficulty. Nor does it depend on any metaphysical theses as contentious as Hume's dictum. It does not require its exponents either to accept or to reject Hume's dictum. Nor does it require them to accept or to reject the existence of powers. Nor, finally, does it require them to accept or to reject something like Lewis' extreme modal realism. It is neutral on all these metaphysical points. If definitions in the philosophy of mind are better for being less metaphysically contentious, then the definition proposed in Section 11.2 would appear to be better than supervenience-/necessitation-based definitions that presuppose abundant properties.

11.7 Conclusion

I've formulated some working definitions of mind-body theories. The most important of these for our purposes is the definition of physicalism. I've defined it as the claim that everything can be exhaustively described and explained by the most empirically adequate theories of current or future physics, and have defended this definition against a variety of objections. I've argued that this definition is superior to definitions that take physicalism to be compatible with property dualism. These include definitions that take nonreductive physicalism to imply a commitment to property dualism. I've also argued that this definition is superior to definitions that are based on notions of supervenience or necessitation. These definitions have difficulty accommodating the core physicalist thesis. In addition, they typically presuppose abundant properties, and a commitment to abundant properties carries with it a number of undesirably contentious metaphysical assumptions. With the foregoing points in mind, we are in a position to consider Williams' worry, the suspicion that hylomorphism is really a form of physicalism in disguise.

12

Williams' Worry

Is Hylomorphism Just a Form of Physicalism?

12.1 Why Hylomorphism Is Incompatible with Physicalism

Chapter 10 introduced Williams' worry, the suspicion that hylomorphism is just a form of physicalism. We saw in Chapter 11 that physicalism can be understood as the claim that everything can be exhaustively described and explained by the most empirically adequate theories of current or future physics. Hylomorphism rejects this claim because according to hylomorphists, there are structures which the conceptual resources of physics by themselves do not enable us to describe and explain. Describing and explaining those structures, they claim, requires the additional conceptual resources of special sciences such as biology and psychology. The claim that there are such structures is largely an empirical one on the hylomorphic view. We saw earlier that biologists, neuroscientists, and others posit structure or organization as an ontological and explanatory principle. According to ontological naturalism, which was adopted as a working assumption in Section 1.2, these empirical sources give us good prima facie reason to think that there are distinctively biological or psychological structures, ones that cannot be described using the conceptual resources of physics alone. This claim is open to empirical falsification. Perhaps hylomorphists are wrong about biological or psychological structures. Perhaps those structures can ultimately be identified with complex relations that are exhaustively describable by physics. Or perhaps there are no structures at all; perhaps structure realism in general is false. I address these issues in greater detail in Chapter 14, but at present, our question is not whether hylomorphism is true, but whether hylomorphism implies physicalism. To answer this question, we need to assume for the sake of argument that hylomorphism is true and show that physicalism can still be false.

We agreed in Section 1.3 that we would use the term 'hylomorphism' to refer to a structure realist view that claims that there are structures other than those postulated by physics. We agreed, moreover, that according to hylomorphism, these structures are basic: they cannot be identified with structures postulated by physics. Suppose, then, that hylomorphism is true, and that there are basic structures, such as biological ones, that cannot be exhaustively described by physics. Since physics cannot

exhaustively describe these structures, and since these structures cannot be identified with structures that can be exhaustively described by physics, it follows that physics cannot exhaustively describe and explain everything. Physicalism, however, is precisely the claim that physics can exhaustively describe and explain everything. Consequently, if hylomorphism is true, physicalism must be false.

12.2 Worry 1: Biological and Physical Structure

The foregoing considerations put us in a position to address Williams' Worry 1. According to Worry 1, recall, hylomorphism must be committed to physicalism because biological structures are just complex relations among physical materials whose properties and behavior can be exhaustively described by physics. Hylomorphists respond by denying the worry's premise: biological structures are not the kinds of relations that physics can exhaustively describe. It is true, according to hylomorphists, that biological structures supervene on properties or relations that can be exhaustively described by physics. We saw that this is an implication of structo-physical necessitation (Section 9.3). For any given higher-level structure such as the structure that makes Godehard what he is, or the structure that makes my throwing a baseball at *t* what it is, there are complex properties or relations among the materials composing and surrounding the individual or event. Many of these properties or relations can be exhaustively described by the most empirically adequate theories in current or future physics, but according to hylomorphists, none of these properties or relations is the kind of dynamic higher-level structure that belongs to the subject matter of biology, psychology, or other special sciences.

Again, it is possible that hylomorphists could be wrong about this; it is an open empirical possibility that the biological, psychological, and other higher-level structures that hylomorphists believe cannot be identified with structures described by physics will turn out to be thus identifiable. Like physicalists, hylomorphists are committed to an empirical conjecture: their theory functions as a "high-level empirical hypothesis," to use Hartry Field's (1972: 357) term. But if hylomorphism is true, then the biological, psychological, and other high-level structures postulated by the special sciences are not identical to complex properties or relations that can be exhaustively described by physics. This neutralizes Worry 1.

12.3 Worry 2: Exhaustive Physical Decomposition

Hylomorphism is committed to the claim that structured individuals and their activities are exhaustively decomposable into the activities and subactivities of their parts and surrounding materials, and to these parts and materials being in turn exhaustively decomposable into fundamental physical materials. Let us call this claim *exhaustive physical decomposition*. According to Worry 2, exhaustive physical decomposition is tantamount to physicalism. Consequently, because hylomorphism

is committed to exhaustive physical decomposition, it must be committed to physicalism as well.

In response, hylomorphists challenge the premise that exhaustive physical decomposition implies physicalism. Physicalism is the thesis that everything can be exhaustively described and explained by the most empirically adequate theories in current or future physics, and this thesis is not implied by exhaustive physical decomposition. It is possible for individuals that are exhaustively decomposable into fundamental physical materials to have first-order nonphysical properties, and for their behavior to be governed by emergent laws in addition to those governing their fundamental physical constituents. This was a key feature of the organismic dual-attribute theories (DATs) described in Section 11.2. We used those theories as a litmus test for definitions of physicalism. Any adequate definition of physicalism, we said, had to imply the core physicalist thesis, the thesis that everything is physical. Definitions of physicalism that failed the test included definitions that were compatible with DAT, including the definition that persons are physical. The claim that persons are exhaustively decomposable into fundamental physical materials is analogous. Even if persons are exhaustively decomposable into fundamental physical materials, and their activities are decomposable into the activities of fundamental physical materials, this does not imply that everything is physical since it is still possible that persons or their activities might have nonphysical properties.

Exhaustive physical decomposition would imply that all properties are physical if it were combined with a thesis like the following:

> *Property exhaustion thesis*: Necessarily, for any x, if x is exhaustively decomposable into $y_1, y_2, \ldots, y_n$, and the activities of x are exhaustively decomposable into the manifestations of the powers of the ys and surrounding materials, then x has no properties other than those of the ys and the surrounding materials.

We have seen, however, that hylomorphists reject any such thesis. They claim that some of a composite individual's properties depend on its structure (Section 6.3). Even if x is exhaustively decomposable into the ys, x will still have some properties that the ys lack, namely the properties due to its structure. The same is true mutatis mutandis of x's activities on the hylomorphic view. If my throwing a baseball is exhaustively decomposable into the manifestations of the powers of my parts and surrounding materials, there is still more to my throwing a baseball than that: there is the way those manifestations are structured or coordinated. In order for me to throw a baseball, the powers of my parts and surrounding materials cannot manifest themselves in just any old way; the right kind of coordination is essential to them composing the activity they do (Sections 8.1–8.2). That coordination, however, is not a property of the parts and materials or their powers taken on their own; it is rather a property of the individual as a whole—a structure the individual imposes on them. Hylomorphism thus implies that anything like the property exhaustion thesis is false.

But unless something like the property exhaustion thesis is true, exhaustive physical decomposition does not imply physicalism. Worry 2 is thus based on a false premise.

12.4 Worries 3 and 4: Necessitation, Physicalism, and Reduction

Hylomorphism implies robust necessitation and supervenience theses (Chapter 9). Worry 3 claims that necessitation and/or strong supervenience is sufficient for physicalism, and hence hylomorphism must be a form of physicalism. We saw in Chapter 11, however, that the Worry's premise is false: neither necessitation nor strong supervenience is sufficient for physicalism. Definitions based solely on necessitation or strong supervenience fail to imply the core physicalist thesis without recourse to contentious metaphysical assumptions. These include the assumption that properties are abundant and the assumption that there can be no metaphysically necessary connections between distinct properties (Sections 11.5–11.6). Hylomorphists reject both of these assumptions. The hylomorphic metaphysic developed in Chapters 2–10 claims that properties are sparse, not abundant. It also claims that properties are powers, and this implies that there are metaphysically necessary connections between distinct properties. In particular, powers are distinct from their manifestations since they can exist without ever being manifested, yet the connections between powers and their manifestations are metaphysically necessary: a fragile object must break if struck the right way; it is metaphysically impossible for something to be fragile and not breakable. So according to hylomorphists, the assumptions on which Worry 3 is based are false. Unless those assumptions are true, necessitation and supervenience fail to imply physicalism. This resolves Worry 3.

Worry 4 claims that necessitation and/or strong supervenience is sufficient not for physicalism, but for reduction, and hence hylomorphism must be committed to reductivism: it must claim that thoughts, feelings, perceptions, and intentional actions are reducible to descriptions and explanations that appeal to physiological mechanisms alone. We saw in Chapter 11 that Worry 4's premise is false. Reduction involves descriptive and explanatory takeover, and this requires that the entities postulated by the reduced theory or conceptual framework be identified with the entities postulated by the reducing theory or conceptual framework (Section 11.3). But neither necessitation nor supervenience implies that the necessitated or supervening properties are identical to the properties by which they are necessitated or on which they supervene.[1] Consider again the kind of view that Stoljar (2010) calls 'necessitation dualism' (Section 11.6). It claims that there are physical properties and distinct nonphysical mental properties. These properties, it says, are connected by

[1] McLaughlin (1995: 45–8) argues for this thesis on different grounds.

metaphysically necessary psychophysical laws. In all metaphysically possible worlds, when something has a nonphysical mental property, M, it has a specific physical property, P, and in all metaphysically possible worlds anything that has P has M. According to the theory, therefore, mental properties are necessitated by physical properties, and since necessitation implies strong supervenience, mental properties strongly supervene on physical properties as well. Yet, the theory rejects physicalism, for according to it, not everything is physical; in particular, it claims that there are nonphysical mental properties. Consequently, neither psychophysical necessitation nor psychophysical supervenience implies that mental properties are identical to physical properties, and hence neither implies a commitment to psychophysical reductivism. What is true of the psychophysical case, moreover, is generalizable to all cases: neither the necessitation nor the supervenience of higher-level phenomena on lower-level phenomena is sufficient for reduction.

Defenders of Worry 4 might argue that there can be no laws of the sort that the foregoing necessitation dualism envisions: no metaphysically necessary connections between distinct properties—a claim called Hume's dictum. We saw in Section 11.6, however, that there are good reasons to think Hume's dictum is false, and it is in any event a claim hylomorphists reject. So, if hylomorphism is true, neither necessitation nor strong supervenience is sufficient for reduction. Consequently, the hylomorphic commitment to necessitation and strong supervenience does not imply a commitment to reductivism. This dispenses with Worry 4.

12.5 Worry 5: Phenomenal Consciousness

Hylomorphism claims that thought, feeling, and perception are composed of the structured manifestations of the powers of lower-level things, and these things, their powers and manifestations, are typically revealed through functional analysis (Sections 8.5, 10.3–10.4). According to Worry 5, this suggests a materialist solution to what Chalmers calls the *hard problem of consciousness*:

> The word "consciousness" is . . . sometimes used for the ability to discriminate stimuli, or to report information, or to monitor internal states, or to control behavior . . . [T]hese phenomena [pose] the "easy problems" of consciousness . . . There seems to be no deep problem in principle with the idea that a physical system could be "conscious" in these senses . . . For these problems, the task is to explain certain behavioral or cognitive functions: that is, to explain how some causal role is played in the cognitive system . . . To explain the performance of such a function, one need only specify a mechanism that plays the relevant role. And there is good reason to believe that neural or computational mechanisms can play those roles. What makes the hard problem hard? Here, the task is not to explain behavioral and cognitive functions: even once one has an explanation of all the relevant functions in the vicinity of consciousness . . . there may still remain a further question: why is the performance of these functions accompanied by experience . . . A mental state is conscious when there is something it is like to be in that state. Conscious states include states of perceptual experience, bodily sensation, mental imagery,

emotional experience, occurrent thought, and more . . . Each of these states has a *phenomenal character*, with *phenomenal properties* (or *qualia*) characterizing what it is like to be in the state . . . There is no question that experience is closely associated with physical processes in systems such as brains . . . But how and why do physical processes give rise to experience? Why do not these processes take place "in the dark," without any accompanying states of experience? This is the central mystery of consciousness. (2002a: 247–8)

What is characteristic of materialist solutions to the hard problem, says Chalmers, is that they see consciousness as a physical process (2002a: 248). That seems to be the way hylomorphists see it, says Worry 5. According to hylomorphists, we develop a mechanistic understanding of ourselves and our powers (including our powers to think, feel, and perceive) in terms of the powers of our parts and surrounding materials. Those parts and materials, and the powers they have, are ultimately all physical. Consequently, if 'consciousness' refers to anything in a hylomorphic framework, it must refer to something that can be understood in physical terms.

There are two arguments in the vicinity of Worry 5. The first runs as follows: (1) hylomorphism is committed to providing mechanistic explanations of our powers to think, feel, and perceive, and that would seem to include the power to have phenomenally conscious experiences. But it seems that (2) mechanistic explanation is the same as what Chalmers calls *reductive explanation*. So it looks like hylomorphism is committed to providing a reductive explanation of conscious experiences. But according to Chalmers, (3) if higher-level phenomena are reductively explainable by lower-level phenomena, then higher-level phenomena logically supervene on lower-level phenomena. (4) Logical supervenience is sufficient for materialism. So it looks like hylomorphism is committed to materialism. But (5) 'materialism' is synonymous with 'physicalism.' So it looks like hylomorphism must be committed to physicalism.

Hylomorphists can respond to this argument by challenging Premise (5). Although 'materialism' and 'physicalism' are often both used as labels for the thesis that everything is physical, and can thus be used interchangeably in some contexts, it doesn't follow that they can be used interchangeably in all contexts. Some authors use 'materialism' and 'physicalism' to refer to different theses.[2] Other authors take them to be synonymous, but do not take physicalism/materialism to imply the core physicalist thesis that everything is physical. We saw several examples of this in Chapter 11. Finally, some authors take physicalism/materialism to imply the core physicalist thesis, but offer definitions that fall short of implying the core physicalist thesis in fact. Chalmers is an example. He defines physicalism/materialism as follows:

The widely held doctrine of *materialism* (or *physicalism*) . . . is generally taken to hold that everything in the world is physical, or that there is nothing over and above the physical . . .

[2] Fodor (1974: 100ff.) is an example.

[M]aterialism is true if all the positive facts about the world are globally logically supervenient on the physical facts. (1996: 41)

Pace what Chalmers says, we've seen that supervenience-based definitions like his have difficulty capturing the core physicalist thesis (Section 11.6).[3] From this it follows that a commitment to materialism in Chalmers' sense does not by itself imply a commitment to physicalism.

The foregoing considerations are sufficient to dispense with the first argument in the vicinity of Worry 5. But it is possible for hylomorphists to say more. Premise (2) of the argument claims that mechanistic explanation is the same as what Chalmers calls reductive explanation, but it is not evident that this is the case. A reductive explanation in Chalmers' sense has two components: it provides an analysis of an explanandum phenomenon in terms of the performance of various functions, and it supplies a list of lower-level items that either do perform or could in principle perform those functions (Chalmers 1996: 42–7). Chalmers makes it clear that by 'function,' he has in mind something like a causal role: the notion of function, he says, "is taken causally rather than teleologically" (44; cf. 105). By contrast, according to hylomorphists, mechanistic explanations postulate components that contribute teleologically to an activity as a whole (Section 6.4).[4] In addition, mechanistic explanations are answers to how-questions of mechanism (Section 10.3), but it is less clear what kinds of explanations Chalmers' reductive explanations are supposed to be. He does not situate reductive explanations in his sense among a broader constellation of explanation types like those described by van Fraassen (1980) and Jaworski (2009). He only characterizes them as mystery-removing explanations, in contrast to illuminating ones:

[A] reductive explanation is not necessarily an *illuminating* explanation. Rather, a reductive explanation is a *mystery-removing* explanation . . . Its chief role is to remove any deep sense of mystery surrounding a high-level phenomenon . . . Once we have told the lower-level story in enough detail, any sense of fundamental mystery goes away: the phenomena that needed to be explained have been explained . . . [A] reductive explanation . . . eliminates any sense that there is something "extra" going on. (Chalmers 1996: 42, 48–9)

Hylomorphists are not committed to mechanistic explanations playing a mystery-removing role. The purpose of mechanistic explanations is to answer how-questions of mechanism; whether they play some further mystery-removing role is incidental. These considerations provide some reason for rejecting Premise (2) in addition to rejecting Premise (5).

[3] Jessica Wilson (2005) argues against supervenience-based definitions of physicalism on independent grounds.

[4] Since this is one of the ideas that hylomorphism has in common with teleological functionalist theories like Lycan's (1987), it is worth considering whether the latter theories count as materialist in Chalmers' sense as well; they very well may not.

There is nevertheless a second argument in the vicinity of Worry 5. To motivate it, let us grant Premise (2) of the first argument: let us assume that mechanistic explanations correspond to reductive explanations in Chalmers' sense. According to the second argument, even if a hylomorphic approach to phenomenal consciousness is not sufficient to establish that hylomorphism is a form of physicalism, it might still be sufficient to establish that hylomorphism is false. Premises (1)–(4) of the first argument imply that if hylomorphism is true, materialism in Chalmers' sense must be true as well. But Chalmers has an argument to the effect that materialism in his sense must be false. Consequently, Chalmers' argument would establish that hylomorphism is false. To address this argument, we will have to consider Chalmers' argument against materialism in detail.

12.6 Chalmers' Argument against Materialism

Chalmers' (2009; 2010: chapter 6) most direct argument against materialism runs as follows:

(1) P & ~Q is ideally 1-conceivable.
(2) If P & ~Q is ideally 1-conceivable, then P & ~Q is 1-possible.
(3) If P & ~Q is 1-possible, then P & ~Q is 2-possible.
(4) If P & ~Q is 2-possible, then materialism is false.

Therefore, materialism is false.[5]

Here P is a true description formulated in the language of microphysics of the fundamental properties of every fundamental microphysical entity, and Q is a truth about phenomenal consciousness—the truth that, say, a particular individual is in some phenomenally conscious state.

[5] Chalmers' fully developed version of the argument includes a technical wrinkle that needn't concern us here. That wrinkle requires him to add a disjunct about Russellian monism to Premise (3): If P & ~Q is 1-possible, then P & ~Q is 2-possible or Russellian monism is true. The argument thus concludes that either materialism is false or Russellian monism is true. Russellian monism is the view that the intrinsic properties of fundamental entities are categorical bases both for the dispositional properties of those entities and for whatever phenomenal properties there are. If Russellian monism is true, it is possible that fundamental physical entities in the worlds w_1 and w_2 might have all the same dispositional properties while yet having different intrinsic properties. We might imagine, for instance, that in w_1 fundamental physical entities have only *F*-properties intrinsically, while in w_2 they have only *G*-properties intrinsically. Still, if Russellian monism is true, it is possible that *F*-properties in w_1 and *G*-properties in w_2 might underwrite the same dispositional properties. According to Russellian monism, the latter properties are precisely those that fix the referents of terms in P, the true physical description of the fundamental (dispositional) properties of every fundamental entity. Consequently, if Russellian monism is true, it is possible that w_1 and w_2 might be physically indistinguishable from each other even though the intrinsic properties of fundamental entities in w_1 are entirely different from the intrinsic properties of fundamental entities in w_2. (The situation is similar to the one Lewis (2008) describes when defending Humility (Section 5.2).) According to Russellian monism, however, phenomenal properties are grounded in the intrinsic properties of fundamental entities, not in their physical dispositions. So even though w_1 and w_2 might be physically indistinguishable, it is still possible that one of them might have a phenomenal truth that the other lacks. Russellian monism is thus compatible with the possibility that P & ~Q.

According to Premise (1), it is conceivable that the world could be physically exactly as it is in fact and yet not include a particular phenomenal truth. It is conceivable, for instance, that a world physically indistinguishable from the actual world might include physical duplicates of us who have qualitatively different conscious states or who lack conscious states altogether. The kind of conceivability Chalmers has in mind is ideal 1-conceivability. To understand what this is, we first have to understand something about the general two-dimensional semantics he endorses.

According to Chalmers, terms have both primary and secondary intensions. Intensions are functions from possible worlds to extensions. A term's primary intension, according to Chalmers, assigns to it an extension comprising whatever has the kinds of characteristics a speaker would use to fix the referent of that term in the world the speaker inhabits. For example, the primary intension of 'water' assigns to it an extension comprising (roughly) all the clear, colorless, odorless, drinkable liquid that fills rivers and oceans in the speaker's world. Consequently, if that liquid is H_2O, as it is in the actual world, then the primary intension of 'water' assigns the extension H_2O: this is what 'water' refers to in that world. If, on the other hand, the clear, colorless, odorless, drinkable liquid that fills rivers and oceans in the speaker's world is XYZ, then the primary intension of 'water' assigns the extension XYZ: XYZ is what 'water' refers to in that world. A term's secondary intension, by contrast, assigns an extension in all counterfactual worlds comprising whatever satisfies the term's primary intension when the speaker's world is treated as actual. In our world, for instance, the secondary intension of 'water' assigns the extension H_2O in all counterfactual worlds. If, however, we inhabited world *w* in which the clear, colorless, odorless, thirst-quenching liquid that fills rivers and oceans is not H_2O, but XYZ—if, in other words, *w* had turned out to be the actual world—then the secondary intension of 'water' would have assigned to the term 'water' not the extension H_2O, but the extension XYZ.

Primary intensions are what interest Chalmers most. The reason is that determining the primary intensions of terms is ostensibly something that can be done a priori. Determining the primary intension of 'water' does not require empirical investigation into what the clear, colorless, odorless, drinkable liquid that fills rivers and oceans actually is; intuitively, it requires only armchair reflection on the term's meaning, in the sense of what a speaker would pick out with the term in the world he or she inhabits (Chalmers 2004). To determine the primary intension of 'water,' I need only reflect on what stuff in a world a speaker would refer to using the term 'water'—roughly, all the clear, colorless, odorless, drinkable liquid in the speaker's surrounding environment.

The apriority of primary intensions is important because according to Chalmers, it licenses inferences from conceivability to possibility. In particular, ideal 1-conceivability is supposed to be a guide to 1-possibility. 1-conceivability is the kind of conceivability that corresponds to primary intensions. It is 1-conceivable,

for instance, that water should not be H_2O, for we can conceive of situations in which the primary intension of 'water' assigns to it a referent other than H_2O. By contrast, it is not 2-conceivable that water should not be H_2O. We cannot conceive of a situation in which the secondary intension of 'water' assigns to it a referent other than H_2O, for that intension assigns to 'water' the referent H_2O in all counterfactual worlds, given that 'water' refers to H_2O in the actual world. If the terms in the statement 'water is not H_2O' are understood according to their secondary intensions, then it is not 2-possible that water should fail to be H_2O: there is no metaphysically possible world in which water fails to be H_2O. It is, however, 1-possible that water should fail to be H_2O, and we can see this, says Chalmers, if we take the terms in 'water is not H_2O' according to their primary intensions. Clearly, there are worlds in which the primary intension of 'water' assigns to it a referent other than H_2O: these are worlds in which the clear, colorless, odorless, drinkable liquid that fills rivers and oceans is not H_2O, but XYZ. 1-conceivability is thus an indicator of 1-possibility, but not just any kind of 1-conceivability will do. It must be ideal conceivability, not mere prima facie conceivability: it must be the kind of conceivability that would be achieved on ideal rational reflection, not merely on initial consideration. According to Premise (2) of Chalmers' argument, it is this kind of conceivability that is sufficient for 1-possibility.

Suppose, then, that it is 1-possible for the world to be physically exactly as it is in fact and yet not to include a particular phenomenal truth. According to Premise (3), this implies that it is 2-possible for the world to be physically exactly as it is in fact and yet not to include a particular phenomenal truth. The argument for Premise (3) relies on the claim that the primary and secondary intensions of P and also of Q assign the same extensions. The case of 'water' is one in which the primary and secondary intension assign different extensions. It is possible for something to resemble water in all its reference-fixing respects without being water. But according to Premise (3), the cases of P and Q are different. Consider first Q. According to Chalmers, it is not possible for something to resemble consciousness in all the respects that fix the referent of 'consciousness' without being consciousness. It is thus plausible to suppose that the primary and secondary intensions of a Q-term such as 'consciousness' assign the same extension. The case of terms in P is somewhat harder to make out, says Chalmers, for materialists might plausibly claim that the primary intension of a P-term such as 'mass' picks out whatever it is in a world that plays a certain theoretical role, while the secondary intension picks out whatever it is that plays that role in the actual world. If it is possible for different properties to play the mass role, then the primary and secondary intensions of 'mass' will not assign the same extensions, and as a result, 1-possibility will not imply 2-possibility. Yet, Chalmers has an easier time making the case against hylomorphists, for as we've seen, they are committed to saying that a property's theoretical role is essential to it (Section 4.4). They will thus deny that different properties can play the mass role in different possible worlds, and will therefore be committed to something like Premise (3).

Finally, Premise (4) claims that if P & ~Q is 2-possible, then materialism must be false. Recall that by 'materialism,' here we mean not just physicalism, but any view that claims that Q logically supervenes on P. We agreed that hylomorphism belongs in this category because it is committed to strong necessitation and supervenience theses. From Chalmers' premises (1)–(4), it follows that materialism (including ostensibly hylomorphism) must be false. Hylomorphists are thus obliged to reject one of Chalmers' premises. In what follows, I'll try to sketch a general line of response. I'll start by explaining the gist of it and will then try to formulate it in terms of Chalmers' preferred categories. The points it covers are not new.[6] If there is any novelty in the response I articulate, it is in the way it situates some well-rehearsed points within a hylomorphic framework.

The way hylomorphists respond to Chalmers' argument is similar in its outlines to the way they respond to other conceivability-based arguments (Sections 4.4, 6.1). Hylomorphists look to reject either Premise (1) of Chalmers' argument or Premise (2), depending on how 'conceivability' and related terms are defined. If we use Chalmers' preferred definitions, then the hylomorphic response is best understood as rejecting (2), the claim that conceivability implies possibility. Our conceptions of things, say hylomorphists, needn't correspond to genuine possibilities. The reason is that our concepts needn't correspond to genuine properties. Those properties, however, constrain the space of metaphysically possible worlds since the metaphysics of powers that hylomorphists endorse implies that natural necessity is a species of metaphysical necessity (Section 4.4). Determining which concepts do and which do not correspond to genuine properties is largely an empirical undertaking. Consequently, if we deploy our concepts a priori even under ideal circumstances, the conceptions we form can fail to disclose what is possible.

In terms of Chalmers' preferred categories, the hylomorphic response rejects the claim that every conceivable scenario—every maximal a priori coherent hypothesis—corresponds to a metaphysically possible world (2009: 326–7). According to Chalmers, a response along these lines is committed to strong a posteriori necessities—'strong necessities,' for short:

> [A]n a posteriori necessity [S] is a . . . *strong necessity*, iff S has a necessary primary intension . . . What would a strong necessity involve? . . . One could put the matter by saying that there is an epistemically possible *scenario* verifying [~S], but no metaphysically possible world verifying [~S]. Here a scenario can be understood as corresponding to a maximal a priori coherent hypothesis, in the way that worlds correspond to maximal metaphysically possible hypotheses . . . When S is a strong necessity . . . there will be a scenario verifying ~S, but this scenario will correspond to no metaphysically possible world . . . [T]he space of . . . metaphysically possible worlds is *smaller* than the space of epistemically possible scenarios . . . [T]here are scenarios that correspond to no worlds . . . [T]here will be a *scenario* verifying P & ~Q, including various

[6] Earlier exponents of it include Peacocke (1989: 67–9), Loar (1990: 82–95), and Papineau (1993: 111–14).

specific zombie scenarios. But these scenarios will correspond to no metaphysically possible world. (2009: 325–7)

Chalmers considers several views which purport that there are strong necessities. The one he attributes to Sydney Shoemaker corresponds most closely to the hylomorphic view:

[L]aws of nature are not just naturally necessary but are metaphysically necessary... [I]f mass in the actual world obeys certain laws, then nothing in any counterfactual world counts as mass unless it obeys exactly those laws, so any law involving mass will be necessary... [P]roperties such as mass have their nomic profile essentially, as on Shoemaker's view... [T]he fundamental properties and laws of all worlds are the fundamental properties and laws of our world (and... these laws are not knowable a priori). In effect, this restricts the space of metaphysically possible worlds to the space of naturally possible worlds. If this view is correct, then a fundamental law will be a strong necessity: there will be no world corresponding to the scenario that we conceive when we conceive what is false. (2009: 329–30)

Chalmers goes on to claim that there is "no good reason" to endorse this kind of view (2009: 330). But the considerations advanced in earlier chapters suggest that this claim is inaccurate. First, competing categoricalist views, according to which laws of nature are metaphysically contingent, face serious difficulties (Chapter 5). In addition, there are good reasons to think that properties are sparse (Chapter 2), that the only properties that exist are ones that make a causal difference to their bearers. There are also good reasons to think that sparse properties are powers, and if sparse properties are powers, then natural necessity is a species of metaphysical necessity (Chapter 4). Chalmers' claim that there is no good reason to endorse a view that posits metaphysically necessary natural laws would thus appear to be false.

But if there are metaphysically necessary natural laws, then it is possible for there to be strong necessities. We saw in Chapter 9 that according to hylomorphists, these could include hylomorphic necessitation relations, which are metaphysically necessary a posteriori relations between an individual's activities and the powers of its parts. If there are strong necessities of this sort, then Premise (2) of Chalmers' argument is false.

12.7 Strong Necessities

There is one more argument that hylomorphists need to address to deal with Chalmers' argument against materialism. It concerns strong necessities. Chalmers (1996: 136–8) advances six reasons for rejecting them:

(i) strong necessities cannot be supported by analogy with other a posteriori necessities; (ii) they involve a far more radical sort of a posteriori necessity than Kripke's, requiring a distinction between conceptual and metaphysical possibility at the level of worlds; (iii) they lead to an ad hoc proliferation of modalities; (iv) they raise deep questions of coherence;

(v) they will be brute and inexplicable; and (vi) the only motivation to posit a strong necessity in the mind-body case is the desire to save materialism. (2009: 327)

The first point to make in response to Chalmers is that the strong necessities hylomorphism postulates are not brute and inexplicable, as (v) claims. They are instead implied by more fundamental hylomorphic principles. According to hylomorphists, we are essentially composed of physical materials with structures (Sections 6.1–6.3), and these structures have metaphysically necessary lower-level material conditions (Section 10.4). It is a brute fact on the hylomorphic view that structures are things of this sort, but the strong necessities hylomorphism postulates are not themselves brute facts; they are instead implications of the more basic hylomorphic principles I've just described. Claim (v) is therefore false. Something analogous can be said in response to claim (vi): strong necessities in the mind-body case follow from basic hylomorphic principles (Chapter 9), they are not introduced simply to snatch hylomorphism from the jaws of arguments like Chalmers'.

The other considerations Chalmers advances against strong necessities center around claim (ii). According to Chalmers, strong necessities require us to distinguish conceptual from metaphysical possibility. If this were the case, then worries (iii) and (iv) would likely follow. But according to hylomorphists, their own postulation of strong necessities does not depend on distinguishing conceptual from metaphysical necessity. Rather, they say, some of our conceptions correspond to no possibilities at all—neither metaphysical possibilities nor possibilities of other sorts. To understand this response, it will be helpful to consider some analogous responses.

The first analogous response is advanced against Descartes by Antoine Arnauld, author of the *Fourth Set of Objections* to the *Meditations*. Descartes argues that we can exist without bodies because thought is our only essential property. To support this premise, Descartes compiles a list of the properties that people initially take themselves to have (extension, color, movement, nutrition, sensation, and so on), and argues that we can clearly and distinctly conceive of ourselves existing without each of the properties on the list—each, that is, except for thought. We cannot form a (first-personal) conception of ourselves without thought, Descartes says, so he claims that we cannot exist without it. Since this is the only property of which this is true, Descartes concludes that it must be our only essential property.

Arnauld's worry about Descartes' argument is that the initial conception we have of ourselves might be in some way impoverished: we might have properties, perhaps even essential ones, of which we are entirely ignorant, and which therefore do not appear on Descartes' original list. We can, he maintains, clearly and distinctly conceive of these properties, and yet even if we do so, Descartes' conclusion will not follow since we might still have essential properties of which we are entirely unaware.[7] Arnauld's point is easy to appreciate. Imagine that object *a* has exactly two

[7] He illustrated the general point with an analogy: "Suppose someone knows for certain that the angle in a semi-circle is a right angle, and hence that the triangle formed by this angle and the diameter of the circle

essential properties: P_2 and P_3. It turns out, however, that the conception we form of a comprises only the properties P_1 and P_2; we know nothing about a's other properties including P_3. Suppose, then, that we follow Descartes' procedure for discerning something's essential properties, and find that we can form a clear and distinct conception of a existing without P_1, but not without P_2. If we take clear and distinct conceivability to be a guide to possibility, we conclude—wrongly—that P_2 is a's only essential property.[8]

Arnauld elaborates this criticism by articulating an account of intellectual abstraction which purports to explain why it is plausible to think that the conception we form of ourselves as thinking things might not comprise all our properties and hence might leave out some essential ones:

> [S]ince I infer my existence from the fact that I am thinking, it is certainly no surprise if the idea that I form by thinking of myself... represents to my mind nothing other than myself as a thinking thing... Hence it seems that this idea cannot provide any evidence that nothing belongs to my essence beyond what is contained in the idea... If you reply that body is not straightforwardly excluded from my essence, but is ruled out only and precisely in so far as I am a thinking thing, it seems that... my knowledge of myself as a thinking thing does not qualify as knowledge of a being of which I have a complete and adequate conception; it seems instead that I conceive it only... by a certain intellectual abstraction. Geometers conceive of a line as a length without breadth, and they conceive of a surface as length and breadth without depth, despite the fact that no length exists without breadth and no breadth without depth. In the same way, someone may perhaps suspect that every thinking thing is also... an extended thing which, besides the attributes it has in common with other extended things... also possesses the peculiar power of thought... [A]lthough... it can by an intellectual abstraction be apprehended as a thinking thing, in reality bodily attributes may belong to this thinking thing. (AT VII, 203–4)[9]

According to Arnauld, we are able to conceive some things in abstraction from others even when they are necessarily conjoined in reality.[10] Even if P_2 and P_3 are both

is right-angled. In spite of this, he may doubt, or not yet have grasped for certain, that the square on the hypotenuse is equal to the squares on the other two sides; indeed he may even deny this if he is misled by some fallacy... [He] clearly and distinctly understand[s] that this triangle is right-angled, without understanding that the square on the hypotenuse is equal to the squares on the other sides. It follows on this reasoning that God, at least, could create a right-angled triangle with the square on its hypotenuse not equal to the square on the other sides... Similarly, although I clearly and distinctly know my nature to be something that thinks, may I, too, not perhaps be wrong in thinking that nothing else belongs to my nature apart from the fact that I am a thinking thing? Perhaps the fact that I am an extended thing may also belong to my nature" (AT VII, 201–3).

[8] Descartes conceded to Arnauld that the conception we started with would have to comprise all our essential properties; it would have to be 'complete', as he put it (AT VII 221–3), but his reasons for asserting that the conception we form of ourselves apart from extension and other corporeal properties is complete in fact remained unclear.

[9] Translations from Descartes' works here and throughout the book are from Cottingham et al. 1984.

[10] Loar has a similar idea: "Does a fully objective description of reality not still leave something out, viz. the subjective conceptions? This is a play on 'leave something out'. A complete objective description leaves out subjective conceptions, not because it cannot fully characterize the properties they discriminate or fully

essential to a, and even if necessarily something has P_2 only if it has P_3, it is still possible for us to form a conception of a that comprises just the former property and not the latter. When we form conceptions in this way, conceivability is not a reliable guide to possibility since the conception we form might exclude properties that are essential, and this could happen even if the conception is clear and distinct, or ideal in Chalmers' sense.

A similar idea about conceivability has been advanced more recently by Eric Marcus (2004). According to him, the claim that zombies are conceivable commits an act-content fallacy; it mistakes the act of conceiving of something without conceiving of its conscious states, on the one hand, for the act of conceiving of something that lacks conscious states, on the other:

> The imagined difference between zombies and us is purely first-personal or subjective . . . a difference in what it's like . . . To imagine zombies first-personally, then, is to imagine what it's like to be a zombie. What is it like to be a zombie? . . . There is *nothing* that it's like to be a zombie. To imagine zombies first-personally is to imagine first-person nothingness . . . So, when we add third-person duplication to first-person absence, we've imagined zombies. But what is it to imagine first-person absence? . . . On the one hand, it is something that we are to imagine from the first-person point of view . . . On the other hand, there is nothing that it's like to be subjectively absent. So there is no imagining of the what-it's-like variety that we can use to arrive at this possibility . . . To 'imagine' creatures that are objectively identical to us with all subjectivity removed is neither an act of third-person imagining, nor an act of first-person imagining. No, to 'imagine' a zombie is not really to imagine at all . . . [T]hose who take themselves to imagine zombies are mistaking not imagining something for imagining nothing. To imagine creatures that are third-personally identical to persons, and first-personally absent is less to imagine than to refrain from imagining, namely to refrain from imagining their inner lives. We imagine creatures just like us, but we diligently refrain from imagining how things are for them . . . No one would argue that to imagine a happy family without imagining their toes is to imagine a toeless happy family. Similarly, it does not follow from the fact that we can imagine creatures third-personally like us without thereby imagining what it's like to be them, that we have imagined creatures third-personally like us whom there's nothing it's like to be. (2004: 482–4)

Other act-content fallacies in the history of philosophy include the one committed by Bishop Berkeley's so-called 'master argument,' which purports to show that the idea of an unconceived object is incoherent (*A Treatise Concerning the Principles of Human Knowledge*, sections 7, 23). According to Berkeley, it is impossible for us to form a conception of an unconceived object. If I try to conceive of an unconceived object, says Berkeley, I will be trying to conceive of something that is both conceived by someone, namely me, and yet not conceived by anyone. But nothing can be both conceived by someone and not conceived by anyone. Hence, Berkeley concludes, the

account for the concepts themselves as psychological states, but simply because it does not employ them" (1990: 93).

idea of a mind-independent object, one that can exist without being conceived by anyone, is incoherent. The principle worry about this argument is that it appears to conflate the act of conceiving with what is conceived. To illustrate this, imagine a parallel argument:

> I am immortal: it is impossible for me not to exist. The reason is that it is impossible to form a conception of myself not existing. Such a conception implies a contradiction because for me to form a conception of myself, I must exist. Hence, for me to form of a conception of myself not existing, I would have to exist, and that means it would be necessary for me both to exist and not to exist. But nothing can both exist and not exist. Hence, it is inconceivable that I should fail to exist. I am thus immortal.

This argument conflates the distinction between my act of conceiving and the situations I am able to conceive. It is true that I cannot engage in the act of conceiving unless I exist, but it does not follow from this that the only situations I am able to conceive are situations in which I exist. I can easily conceive all sorts of situations in which I do not exist—situations that occurred before I was born, for instance. The argument is thus flawed, insofar as it moves illicitly from a premise about my act of conceiving (that I can conceive of something only if I exist) to a conclusion about what I am able to conceive (that I can conceive only of situations in which I exist). Berkeley's master argument commits a similar mistake. It is true that I can conceive of this book only if I am conceiving of it, but it does not follow from this that I can conceive only of situations in which the book is being conceived by someone.

If Marcus is right, something analogous is true of philosophers who claim to be able to conceive of zombies. It is true that I can conceive of humanlike beings who are behaviorally indistinguishable from us without conceiving of their first-personal conscious states, but it does not follow from this that I can conceive of human beings who are behaviorally indistinguishable from us but who lack first-personal conscious states.

Two more arguments about conceptual errors are advanced by Daniel Dennett and Peter Carruthers, respectively:

> Supposing that by an act of stipulative imagination you can remove consciousness while leaving all cognitive systems intact... is like supposing that by an act of stipulative imagination, you can remove health while leaving all bodily functions and powers intact. If you think you can imagine this, it's only because you are confusedly imagining some health-module that might or might not be present in a body. Health isn't that sort of thing, and neither is consciousness. (Dennett 1995: 325)

> Some of our mental-state concepts are *recognitional*... [They] consist in a capacity to recognize, straight off, the corresponding state. I can recognize the feel of pain purely by its feel, without having to appeal to any of my beliefs about causal roles and functions...
>
> Possessing purely recognitional concepts of *feel*, we can deploy those concepts in thought experiments in ways which are unconstrained by the physical or functional facts. But nothing

follows about the non-physical, non-functional nature of the properties which those concepts pick out . . . [If] natural properties can be individuated 'thickly', in independence of any particular way of conceptualizing them, then it is an open question whether the very properties which we recognize on the basis of *feel* may actually be physical . . . [By analogy] suppose that Mary is someone who has been trained to classify chicks into As and Bs—where the As are in fact male, and the Bs are in fact females—but without . . . her having any idea of what it is about the As which underpins recognition. Then we ask Mary: 'Can you conceive of a world which is micro-physically identical with our own, except that the chicks which are As in this world are Bs in that, and vice-versa?' If *A* really does express a purely recognitional concept . . . then she should answer 'Yes'. For then all she has to imagine is that she is confronted with a chick exactly like this A-chick in all micro-physical respects, but that it is one which evokes a recognitional application of the concept *B*. (Carruthers 2000: 55–7)[11]

What Arnauld, Marcus, Dennett, Carruthers, and critics of Berkeley have in common is the idea that there are ways of deploying concepts that are capable of misrepresenting how things are and how they could be. According to Dennett, misrepresentations can arise on account of concepts that are inaccurate or inadequate. According to Carruthers, they can arise on account of deploying concepts independently of one another and illicitly drawing conclusions from this about the possible independence of properties. According to Arnauld, errors can arise in both ways, and according to Marcus and critics of Berkeley, errors can arise from conflating the act of conceiving with the content conceived. I think it is possible to combine the foregoing points into a general response to Chalmers' argument.

We can start by drawing an analogy between conceptual representations and linguistic ones. Roughly, concepts play roles in thought analogous to the roles predicates play in language. If the analogy holds, then what we said in Section 2.2 about the relation between properties and predicates applies mutatis mutandis to the relation between properties and concepts:

- different concepts can correspond to the same property;
- different properties can correspond to the same concept in different contexts;
- not every concept need correspond to a property;
- not every property need correspond to a concept.

These are straightforward implications of the claim that properties are sparse, which was defended in Sections 2.3–2.4. With the distinction between concepts and properties in place, we can identify several ways in which errors can be introduced in the process of conceptualization.

First, our concepts can be inaccurate: we can believe falsely that a concept includes a condition it actually excludes, or that it excludes a condition it actually includes. Someone might believe incorrectly, for instance, that being a bachelor implies being a

[11] Carruthers' discussion is clearly indebted to Loar 1990: 87ff.

woman or being married. If we deploy concepts that are inaccurate, then the conceptions we form needn't correspond to genuine possibilities.

Second, our concepts can be inadequate: we can fail to grasp the full range of conditions for correctly deploying a concept. Someone might know that being a bachelor implies being male but believe incorrectly that being a bachelor has no implications for one's married status. If we deploy concepts that are inadequate, then once again, the conceptions we form needn't correspond to genuine possibilities.

Third, even if we have concepts that are accurate and adequate, we can still fail to deploy them correctly. We might know fully the conditions under which the concepts *BACHELOR* and *MARRIED* are correctly deployed, or the conditions under which the predicates 'is a bachelor' and 'is married' are correctly applied, and yet utter the sentence 'Richard is a married bachelor' or think that Richard is a married bachelor.[12] There are at least two circumstances in which we might do this. First, it could be the case that even though we fully and accurately grasp the conditions for correctly deploying the concepts or applying the predicates, we might fail to appreciate that one of them implies the negation of the other. Roy Sorensen provides an example:

> I long believed that 'The American Thanksgiving Holiday is on the last Thursday of November which is the fourth Thursday in November'. Only in November 2000, which contains five Thursdays, did I realize that these two definite descriptions only partially overlap. Of course, I long knew that November has more than twenty-eight days and that there are only seven days in a week and that the first day of the month cycles forward each year. But I did not pull together all of these analytical truths. (2002: 340)

When we deploy various concepts, we can fail to draw out their implications for each other—fail to pull them together, as Sorensen says.[13] Likewise, we could in principle utter 'Richard is a married bachelor' or think that Richard is a married bachelor on account of failing to draw out the implications that the attribution of bachelorhood has for someone's married status (the example is less credible only because the concepts it involves are less complex). In cases like this, the conceptions we form needn't correspond to genuine possibilities.

In addition, we can alter the force with which we say or think that Richard is a married bachelor. If Richard is a notorious philanderer, for instance, we might say this as a joke. In doing so, we do not intend for our audience to interpret the utterance as a literal assertion of contradiction; we instead intend for them to follow Gricean (Grice 1989) maxims of communication and find a coherent message in what is uttered. We

[12] This presupposes that 'think' is not an achievement verb, contrary to what someone like Gareth Evans (1982) claims.

[13] Something analogous happens when viewing visual paradoxes like Istvan Orósz's *Il Cavallo*, which adorns the jacket of Gendler and Hawthorne 2002. There is nothing paradoxical about the individual elements of the drawing—about the depiction of the second column from the right, or the horse's tail, or the cavalier's sword. Paradoxes arise when we put the various elements of the drawing together and try to see them as depicting a scene that could actually exist.

are able to do so because the conditions for correctly applying the predicate 'is a bachelor' comprise more than simply being unmarried: they also comprise, we might suppose, an association with certain stereotypical bachelor behaviors. This association enables our utterance to retain some cognitive significance, and it is because of this residual cognitive significance that we can reasonably expect an audience to get the joke. We are thus capable of using an expression to achieve a communicative end even if not all the conditions for correctly applying the expression are satisfied in a particular case. In the joke example, we do this by consciously ignoring some of the conditions for correctly applying the predicate 'is a bachelor.'

What we do with predicates we can also do with concepts: we can perform acts of conceiving in which we can put some of the conditions for correctly deploying a concept out of play. Arnauld's geometers put breadth out of play when conceiving of lines, and we put various powers out of play when conceiving of characters like Superman, such as the powers of surrounding materials that would inhibit the kinds of things we imagine Superman doing. When we put conditions for correctly deploying our concepts out of play, the conceptions we form needn't correspond to genuine possibilities.

Someone like Chalmers might insist that when we conceive of characters like Superman, we are conceiving of possible worlds with different laws of nature. Recall, however, that according to hylomorphists, natural necessity is a species of metaphysical necessity (Section 4.4). If this is the case, then there are no possible worlds with different laws of nature. When we imagine characters or scenarios that violate natural laws, we do not succeed in conceiving worlds that are metaphysically possible, nor do we succeed in conceiving worlds that are possible in some other sense—a point to which I'll return momentarily.

Someone might object that in the Superman case we put conditions out of play intentionally: we actually choose to ignore them. But no such thing happens when we conceive of P & ~Q; we do not take ourselves to be putting any of the relevant conditions for correctly deploying our concepts out of play, and so the cases are disanalogous. In response, hylomorphists claim that what we do intentionally we can also do unintentionally. We needn't intend to put conditions for deploying a concept out of play in order to do so, and this is perhaps what happens when some people conceive of zombies or inverted spectra: they put out of play conditions for correctly deploying psychological and/or other concepts not by choice, but unintentionally. We thus form conceptions that do not disclose genuine possibilities.

The cases of inaccurate and inadequate concepts, of failed deployments, and deployments that put conditions out of play involve the content of our acts of conceiving. But if Marcus and critics of Berkeley are right, we can also commit metaconceptual errors in judging what is possible. We can, for instance, mistakenly think that a characteristic of our act of conceiving has implications for the content we conceive. In these cases, we can mistakenly believe ourselves to have formed conceptions of possible things when we have in fact formed no such conceptions at all.

I've just described a number of ways in which we can form conceptions that do not correspond to any metaphysically possible worlds. According to hylomorphists, when people conceive of P & ~Q, they are forming conceptions in one of the foregoing ways. Either (a) they are deploying concepts that are inaccurate or inadequate, or (b) they are deploying concepts incorrectly, (c) they are putting out of play conditions for correctly deploying those concepts, or (d) they are committing a metaconceptual error. Hylomorphists needn't insist that the same kind of error is committed in every case. It is possible for one person to commit error (a) when she conceives of P & ~Q, and for another person to commit error (b) when he conceives of P & ~Q. It is even possible for one and the same person to commit different errors on different occasions. The important point is that hylomorphists have resources for diagnosing the kinds of errors people are likely to make when conceiving of P & ~Q.

With the foregoing points in mind, let us return to Chalmers' claim (ii). The assumption that motivates claim (ii) is that our conceptions must correspond to possible situations—if not metaphysically possible situations, then at least epistemically possible ones, what Chalmers calls 'scenarios.' Chalmers tells us, "My view is that for every scenario there is a corresponding world" (2009: 327). We can call this claim *concept-world correspondence* (CWC). Given CWC, it is easy to see how the postulation of strong necessities would lead to a distinction between metaphysically and conceptually possible worlds: if all our conceptions must correspond to possible worlds, and some of those conceptions correspond to no metaphysically possible worlds, then there must be possible worlds of some other, broader sort to which those conceptions correspond. If we reject CWC, however, this result does not follow. If, moreover, there is no division of worlds into those which are metaphysically possible and those which are possible in some broader sense, then there are no attendant worries about coherence, as (iv) says. Nor is there an ad hoc proliferation of modalities, as (iii) says. What has been said, moreover, should make it evident that the hylomorphic postulation of strong necessities does not depend on drawing analogies with familiar a posteriori necessities; the postulation of those necessities follows instead from more basic hylomorphic principles, so (i) does not appear to be a worry either. The core of the hylomorphic response to Chalmers, then, is to endorse strong necessities and to reject CWC. Hylomorphists claim that there are many ways of deploying concepts that correspond to no possible situations, and if that is the case, then Premise (2) of Chalmers' argument is false. Even if P & ~Q is ideally 1-conceivable, it does not follow that P & ~Q is 1-possible.

Consider now an objection to the hylomorphic response. Someone might object that if we deploy our concepts in any of the foregoing ways, (a)–(d), we are no longer conceiving of something under ideal conditions, and so none of the aforementioned ways of deploying concepts count as ideal conceptions in Chalmers' sense. S is ideally conceivable, according to Chalmers, if S is prima facie conceivable and the justification for S's conceivability cannot be defeated by better reasoning (2002b: 148). In the cases I've described, however, people's conceptions can be defeated by better

reasoning, namely reasoning that avoids errors (a)–(d). As a result, the cases I've described do not constitute counterexamples to Premise (2) of Chalmers' argument, the premise that if P & ~Q is ideally 1-conceivable, then P & ~Q is 1-possible, for in those cases P & ~Q is not ideally 1-conceivable.

In response, hylomorphists say that in cases (a)–(d) it's true that people's conceptions can be defeated by better reasoning, but according to hylomorphists, the better reasoning that defeats those conceptions depends on a posteriori information. Without that extra a posteriori information, the conceptions people form of P & ~Q might be undefeatable, there might be no better conceptions to be formed on the basis of a priori information alone. Critics of hylomorphism might insist that Chalmers' definition of ideal conceivability is meant to rule out defeatability by any better reasoning at all, including reasoning based on a posteriori information. But this cannot be right, for once a posteriori information enters the picture, the result is no longer 1-conceivable. According to Chalmers, 1-conceivability is supposed to be a priori, so a 1-conception has to be based on a priori information alone. But hylomorphists insist that there is no good reason to think that the conceptions people form of P & ~Q can be defeated by a priori information alone. Consequently, they say, P & ~Q can be ideally 1-conceivable and yet neither metaphysically possible nor possible in some broader logical or conceptual sense. When we ideally 1-conceive of P & ~Q, we do not succeed in conceiving any possible world at all.

12.8 Conclusion

The goal of this chapter was to address the first five versions of Williams' worry. Hylomorphism is incompatible with physicalism because it is committed to structures, such as biological and psychological ones, that cannot be described and explained by physics. Worry 1 claims that those structures can be identified with complex relations that can be described and explained by physics, and as a result, hylomorphism is committed to physicalism. Hylomorphists respond by denying that biological, psychological, and other special scientific structures can be identified with complex relations that can be described by physics.

Worry 2 claims that since hylomorphism is committed to individuals and their activities being exhaustively decomposable into fundamental physical materials, it must be committed to some type of physicalist reductivism. Hylomorphists respond once again by denying the Worry's premise. Even if individuals and their activities are exhaustively decomposable into physical materials, it does not follow from this that all the properties of individuals and their activities can be exhaustively described and explained by physics.

Worry 3 claims that hylomorphism must be committed to physicalism because hylomorphism is committed to strong necessitation and supervenience theses. Hylomorphists nevertheless deny that either necessitation or supervenience is sufficient for physicalism. That could be the case only if properties are abundant and there can be no

necessary connections between them. We saw in Chapter 2, however, that there are good reasons to deny that properties are abundant, and we saw in Chapter 11 that there are good reasons to deny that there can be no necessary connections between properties.

Worry 4 assumes that necessitation and/or supervenience are sufficient for reduction, and hence hylomorphism must be committed to some type of reductivism. Once again, hylomorphists respond by denying the Worry's premise. We saw in Chapter 11 that neither necessitation nor supervenience is sufficient for reduction.

There are two arguments in the vicinity of Worry 5. The first argues that hylomorphism must be committed to physicalism because it is committed to giving mechanistic explanations of conscious experiences. These explanations appear to be the same as what Chalmers calls reductive explanations. But a view that claims that it is possible to give reductive explanations of conscious experiences is committed to materialism, and materialism is the same thing as physicalism. Consequently, hylomorphism must be committed to physicalism. Hylomorphists respond that materialism in Chalmers' sense is not the same as physicalism since materialism in Chalmers' sense fails to imply the core physicalist thesis. In addition, hylomorphists argue that most likely mechanistic explanations in their sense are not the same as reductive explanations in Chalmers' sense. So the first argument in the vicinity of Worry 5 has at least one false premise and possibly two.

The second argument in the vicinity of Worry 5 claims that hylomorphism is committed to materialism in Chalmers' sense because it is committed to strong necessitation and supervenience theses, and Chalmers gives us good reason to think that materialism is false. The reason is that it's conceivable that P & ~Q, where P is a true description of the fundamental properties of every fundamental microphysical entity, and Q is a truth about phenomenal consciousness. According to Chalmers, conceivability is a guide to metaphysical possibility, so it's metaphysically possible that P & ~Q. But if that is metaphysically possible, then materialism—including hylomorphism—must be false. Hylomorphists respond by denying that conceivability is in general a guide to metaphysical possibility or to any other kind of possibility. There are many circumstances in which we form conceptions that correspond to no possible situations. These include circumstances in which (a) we deploy concepts that are inaccurate or inadequate, (b) we deploy concepts incorrectly, (c) we put out of play conditions for correctly deploying those concepts, or (d) we commit a metaconceptual error.

Rejecting Chalmers' conceivability-possibility premise requires what Chalmers calls strong necessities—a posteriori truths that have metaphysically necessary primary intensions. Chalmers advances several arguments against strong necessities. Their crux is the concept-world correspondence thesis, the claim that every conceivable scenario corresponds to a possible world. Hylomorphists reject the thesis. Their response thus accepts strong necessities and rejects concept-world correspondence.

The foregoing points address Worries 1–5. Worries 6, 7, and 8 are addressed in Chapter 13, which deals with mind-body problems.

13

Hylomorphism and Mind-Body Problems

13.1 Problems with Emergentism: Emergence and Downward Causation

Chapters 1–10 developed a theory of hylomorphic structure. It is time to consider how that theory handles mind-body problems. I'll start by discussing the problems facing emergentism since they are frequently thought to pose problems for hylomorphism as well. Foremost among them are the problem of emergence and the problem of downward causation.

Emergentism, recall, is a type of dual-attribute theory (Section 11.1). Dual-attribute theories in general are committed to two claims. First, they are committed to property dualism: they claim that there are first-order mental properties and first-order physical properties, and that these mental and physical properties are distinct. Second, they are committed to claiming that the individuals with mental properties have some physical properties essentially. In addition to these two claims, emergentists add two more: mental properties are generated or produced by physical properties, and mental properties are capable of producing physical effects. These additional claims are the basis for the problem of emergence and the problem of downward causation.

The problem of emergence is the problem of explaining how lower-level physical or physiological occurrences can generate or produce higher-level mental phenomena such as consciousness (Jaworski 2011: 229–33). How is it that, say, the movements of tiny particles in my brain can give rise to the rich qualitative experiences I have? Critics argue that emergentists have no satisfactory answer. Some emergentists, for instance, claim that there are psychophysical laws that are every bit as basic as the laws governing purely physical interactions. The world just is a place in which physical events produce mental events in accordance with basic laws. But critics argue that psychophysical laws do not really address the problem. Knowing that there is a psychophysical law linking pain to brain state B might explain why this or that instance of pain is correlated with this or that instance of B, but it does not explain why there is a correlation between instances of pain and instances of B in the first place. Saying that there are psychophysical laws does not explain how the rapid

vibrations of countless tiny physical particles can generate the stable, homogeneous experience of seeing a pink ice cube, it simply says that the generating happens in a regular, lawlike way. Exponents of psychophysical laws take these lawlike correlations to be brute, unexplainable matters of fact. If we want to know how there could be such a law, there is no answer. But according to critics, some further explanation is needed for why there would be laws connecting things as different as brain states and qualitative experiences. Psychophysical laws are every bit as mysterious as the psychophysical correlations they are introduced to explain, and because of that, say critics, they do not offer a satisfactory solution to the problem of psychophysical emergence.

Emergentists are also free to endorse panpsychism: they are free to claim that mental phenomena, such as the conscious experiences you and I have, are present at every level of reality, even the fundamental physical level. If fundamental physical particles have mental states just like ours, then there is no longer a problem explaining how composites of fundamental physical particles such as human organisms could have mental states, for on this view, nothing lacks mental states. Critics nevertheless argue that panpsychism is highly counterintuitive. It is highly counterintuitive that quarks, electrons, and other fundamental physical particles should have mental states like our own. Can we really believe that an electron has a rich qualitative experience of the world the way you and I do?

Emergentists can soften the counterintuitive implications of panpsychism by endorsing panprotopsychism. They can claim that the conscious experiences of fundamental physical particles are not as sophisticated or as rich qualitatively as our own. Strictly speaking, fundamental physical particles do not have mental states of the sort we have; they instead have protomental states—states that are like simpler precursors of the mental states we have. When those particles combine to form atoms or molecules, they give rise to protomental states that are more sophisticated, and when atoms and molecules combine to form neural tissues or different parts of the brain, they give rise to mental states that are more sophisticated still, and so the combinatorial process goes until combinations of lower-level items eventually give rise to the rich, sophisticated mental states we are familiar with in our everyday dealings.

But critics argue that panprotopsychism only solves the problem of emergence by replacing it with a different problem: the problem of saying what protomental states are and how they combine to produce more sophisticated mental states. We know what beliefs, desires, and perceptions are, but what are protobeliefs, protodesires, and protoperceptions supposed to be? I believe that 2+2 = 4, but what would it mean to have a protobelief that 2+2 = 4? Or are protobeliefs not supposed to have propositional contents? If so, then in what sense can they qualify as proto*beliefs*? Panprotopsychists can respond that protomental states are theoretical postulates, that they are precisely the states of lower-level things that in combination produce familiar mental states. But how exactly is any such combination supposed to work?

Panprotopsychism assumes that protomental properties are aggregative properties like mass (Wimsatt 1985). Clearly, if x has a mass of 1kg and y has a mass of 1kg, then x and y have a collective mass of 2kg, but it is unclear how mental properties—or protomental properties—can manage to be aggregative in this way. Unless the foregoing issues are resolved, say critics, panprotopsychism does not really offer a solution to the problem of emergence.

Emergentists can also appeal to the notion of spatial organization or structure. There is more to emergence than physical particles themselves. Not just any spatial arrangement of fundamental physical particles can generate conscious states. A tree trunk and a human brain are composed of the same kinds of fundamental physical particles. Why is it, then, that the brain gives rise to conscious states but the tree trunk does not? The answer, emergentists can say, is that in the brain, but not the tree trunk, those particles are spatially arranged in the right sort of way. But critics insist that this appeal to spatial arrangement does no better than the appeal to brute psychophysical laws. How, after all, is spatial arrangement supposed to explain the emergence of mental states? Suppose a certain number of fundamental physical particles do not generate any mental states (Figure 13.1A). We then rearrange the particles by imposing on them a certain spatial organization (Figure 13.1B). How is this repositioning supposed to explain the emergence of mental states? If the particles in Figure 13.1A do not give rise to consciousness, how could changing their spatial relations make any difference? It does not seem as though it could. According to critics, then, a response that appeals to spatial arrangement does no better than a response that appeals to psychophysical laws alone. Both responses attach a label to a mystery without doing anything to resolve it. It is mysterious how a number of fundamental physical interactions could produce mental states, but it is also mysterious how they could do so in a lawlike way, and how rearranging them could make any difference at all.

The second problem facing emergentists is the problem of downward causation (Jaworski 2011: 238–43). It is the problem of accounting for the causal efficacy of

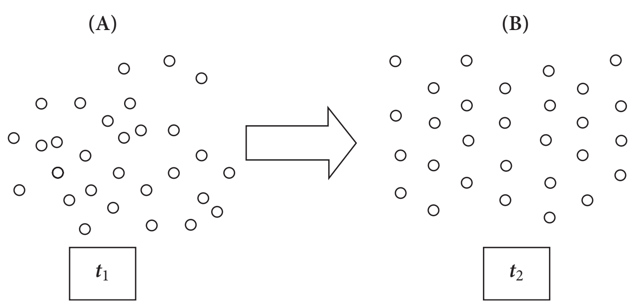


Figure 13.1 Structure and emergence.

higher-level phenomena. The well-rehearsed problem of mental causation is one version of it:

(1) Actions have mental causes.
(2) Actions have physical causes.
(3) The mental and physical causes of actions are distinct.
(4) If actions have multiple causes, then they are overdetermined.
(5) Actions are not overdetermined.[1]

Claims (1)–(3) imply that actions have multiple causes: mental causes and distinct physical causes. By contrast, claims (4)–(5) imply that actions do not have multiple causes. The five claims are thus jointly inconsistent. According to critics, emergentists have difficulty resolving the inconsistency (Kim 2005, 2006b). On the one hand, emergentists cannot reject (1), as eliminative physicalists and epiphenomenalists do. The reason is that emergentists insist that there are thoughts and feelings which are causally efficacious. Given plausible assumptions, moreover, emergentists cannot reject (3), the claim rejected by reductive physicalists. Those assumptions are that (i) causes are events, and that (ii) events are property exemplifications in the sense described in Section 2.1: an event exists exactly if an individual has a property at a time or several individuals stand in a relation at a time, and event e is identical to event e^* exactly if e and e^* comprise the same individuals, properties, and times. Given these existence and identity conditions for events, emergentists must be committed to the distinctness of mental and physical causes. Recall that according to them, mental and physical properties are distinct. Since event identity requires property identity, if mental and physical properties are distinct, then mental and physical events must be distinct. But if mental and physical events are distinct, and causes are events, then mental and physical causes must be distinct. Rejecting (3) is thus not an option for emergentists. Given two further assumptions, moreover, rejecting (4) is not an option either. They are the assumptions that (iii) causes are sufficient conditions, and that (iv) mental and physical causes are independent of each other. Given assumption (iii), if mental event m and physical event p are both causes of action a, then m and p is each sufficient by itself to bring about a. If m and p are independent of each other, moreover, then it looks like a must be overdetermined since an event is overdetermined exactly if it has multiple independent sufficient causes.

That leaves claims (2) and (5): emergentists can either deny that actions have physical causes, or else they can embrace overdetermination. But neither option is attractive. Rejecting claim (2) appears to commit emergentists to saying that physical laws are violated every time we act. We know empirically that our intentional actions are correlated with various physiological events such as events in the central and

[1] Kim's (1993b) exclusion argument is a version of the problem that takes (2) and (5) as nonnegotiable starting points.

peripheral nervous systems. What rejecting (2) seems to imply is that even though these events are occurring, they cease to be efficacious when there is a mental cause at work; that even though neurons continue to fire and muscles continue to contract, these events are not responsible for the action, it is the mental event alone which is. This seems like a rather implausible view. Rejecting (5), on the other hand, implies that actions are overdetermined—that each has both a sufficient physical cause and an independent sufficient mental cause, and this too has awkward implications. In particular, it implies that the action could have occurred without a mental cause or without a physical cause. If certain physiological events were sufficient by themselves to bring about a particular action, then no mental cause was necessary for that action to occur: the action could have occurred without any mental cause at all. If, on the other hand, the mental cause by itself was sufficient to bring about the action, then no physical cause was necessary for the action to occur. The action could have occurred without any physical cause at all. Yet, it seems implausible to suppose that an intentional action could occur without both mental and physical causal factors. It seems implausible to suppose that I could throw a baseball without certain events in my nervous system bringing about my limb movements, and it seems equally implausible to suppose that those limb movements could count as an intentional action without the relevant intentions, beliefs, or desires.[2] On the whole, then, it looks like emergentists have difficulty solving the problem of mental causation in a satisfactory way, and the same is true by extension of the problem of downward causation in general.

13.2 Hylomorphism and the Problem of Emergence

Hylomorphists take mind-body problems to be symptomatic of a worldview that rejects the existence of hylomorphic structure. A worldview that includes hylomorphic structure includes a basic principle that explains why some individuals have distinctive powers. Because structure matters on such a worldview, it carves out distinctive individuals from the otherwise undifferentiated sea of matter and energy that is or will be described by our best physics, and because structure makes a difference, it confers on those individuals distinctive powers, including the powers to think, feel, and perceive. A worldview that does not include hylomorphic structure is one that lacks a basic principle of this sort: it lacks something that distinguishes the parts of the physical universe that can think, feel, and perceive

[2] The neurosurgeon Wilder Penfield's observations support this idea: "When I have caused a conscious patient to move his hand by applying an electrode to the motor cortex of one hemisphere, I have often asked him about it. Invariably, his response was: 'I didn't do that. You did.' When I caused him to vocalize, he said: 'I didn't make that sound. You pulled it out of me.'" (Penfield 1975: 76).

from those that can't. Without a basic principle that carves out zones with distinctive powers, the existence of those powers in the physical universe becomes inexplicable and mysterious. There is nothing that explains why Zone A has powers that Zone B lacks. Nothing explains why you, say, have the power to think, feel, and perceive, while the materials surrounding you do not. Without an explanatory principle of this sort, the options for understanding the existence of those powers in the natural world are constrained. It is possible to claim that their existence is just an inexplicable matter of fact, one that must be accepted with "natural piety," as the emergentist Samuel Alexander (1920: 47) put it. It is also possible to claim that those powers are identical to combinations of the powers of physical materials. Finally, it is possible to deny, as substance dualists and eliminative physicalists do, that those distinctive powers exist in the physical universe at all. But if there is structure, the options are not constrained in this way. The existence of distinctive powers in the natural world is no longer inexplicable. Those powers exist in the natural world because structure exists in the natural world. Nor does this simply push the demand for an explanation back a step, for on the hylomorphic view, structure itself, like material, is a basic principle, one that stands in need of no further explanation.

Consider now the problem of emergence discussed in Section 13.1. It is the problem of explaining how lower-level physical or physiological occurrences can generate or produce higher-level phenomena such as consciousness. How is it that the movements of tiny particles in my brain can give rise to the rich qualitative experiences I have? We saw that various emergentist attempts to answer this question—including the emergentist appeal to spatial organization—face difficulties. Hylomorphists approach the question a different way. They claim that it is illegitimate even to request an explanation of how lower-level factors give rise to higher-level phenomena (Jaworski 2011: 352–3). It is legitimate to request an explanation of how it is possible that *p* only if it is possible that *p*, and according to hylomorphists, it is not possible for lower-level physical or physiological occurrences to produce thoughts, feelings, perceptions, and actions. The reason is that hylomorphism denies that emergent properties are generated or produced by lower-level processes or states.

According to hylomorphists, higher-level phenomena are ways in which lower-level occurrences are structured, and structures in general are not generated or produced by the things they structure. The structure that makes something a piano is not produced by pieces of wood and metal; it is instead something imposed on the wood and metal. Likewise, say hylomorphists, brains do not generate or produce thoughts, feelings, perceptions, and actions. The latter are instead coordinated manifestations of the powers of brains (along with other organic subsystems and surrounding materials). The powers to think, feel, perceive, and act are embodied in muscular contractions, neural firings, and other physiological occurrences, but they are not generated or produced by those occurrences, for on the hylomorphic view,

structured things are not in general causal byproducts of the lower-level things they structure. Requesting an explanation of how lower-level occurrences generate higher-level phenomena thus misunderstands the hylomorphic notion of structure. It assumes, contrary to hylomorphism, that structure is not a basic principle, but is instead something that is derived from lower-level materials. Consequently, demanding that hylomorphists explain how brains produce consciousness implicitly begs the question against their view, for it assumes the existence of a kind of occurrence that hylomorphists deny exists, namely the generation of higher-level phenomena by lower-level occurrences. On hylomorphists' own terms, it is not legitimate to request an explanation of lower-level generation any more than it is legitimate to request that a meteorologist explain how the will of Zeus produces rain. Opponents of hylomorphism are free to reject the view wholesale. But within the hylomorphic framework itself, requesting an explanation of how lower-level phenomena generate higher-level phenomena is illegitimate. Moreover, the very fact that the problem of emergence arises for opponents, but not for hylomorphists, weighs in favor of taking hylomorphic structure as a primitive.

Opponents might argue that hylomorphists face their own problem of emergence. What, after all, explains the emergence of structure itself on the hylomorphic view? Structural phenomena did not always exist in the universe. What, then, were the lower-level physical conditions that were responsible for bringing them about? Hylomorphists respond once again that on their view structure is basic—every bit as basic as the materials that get structured. Asking why either structure or materials exist on the hylomorphic view comes close to asking why the universe as a whole exists, or why there is something instead of nothing. It is possible to reject the hylomorphic worldview and with it the claim that structure is basic. But to request that hylomorphists explain how structure emerges is to request something that hylomorphism implies cannot be done.

Hylomorphists do not deny that we can ask how particular structures came to be in place. It is legitimate to ask how my distinctively human structure came initially to inform various biotic materials. The answer has to do presumably with my parents' reproductive activity. Likewise, it is a legitimate empirical endeavor to attempt to discover how the first living things emerged; that is, how the first living structures came to inform various prebiotic materials. What is not legitimate to ask, according to hylomorphists, is what is responsible for continually generating the structures that I and other living things have. My structure is not something continually generated by some external source or by the materials that compose me; it is instead a self-maintaining configuring activity in which I continuously and essentially engage (Sections 6.1, 6.3). There is no sense, then, in which hylomorphists' refusal to answer a request to explain the emergence of structure can count as a strike against their view, at least not without begging the question and assuming from the outset that the hylomorphic view is false.

There is nevertheless another way of formulating the problem of emergence. It asks what it was in the history of the universe that was originally responsible for the emergence of higher-level phenomena. Biological and psychological phenomena did not always exist in the universe. Early in the universe's history there were not even atoms since energy levels were too high to allow protons and electrons to form stable pairs. Life and mind are thus relative newcomers on the cosmic scene. To many people, this suggests that the physical conditions that existed before their emergence had to be responsible in some way for bringing them about.

It seems increasingly likely that we will be able to discover the physical conditions that were responsible for the emergence of life, but many philosophers are convinced that consciousness is a different story. The following line of reasoning provides one reason to think so: (1) Physical particles are not conscious; moreover, (2) no number of nonconscious particles can combine to produce a conscious whole. After all, one particle by itself does not have the power to produce consciousness. If that were the case, then consciousness would probably have emerged much earlier in the universe's history than it did in fact. In addition, if N particles do not have the power to produce consciousness, then $N+1$ particles do not have the power to produce consciousness, for we just agreed that one particle by itself is not enough to produce consciousness. So if N particles do not produce consciousness, then one more particle is not going to make any difference. Therefore, no number of nonconscious particles has the power to produce consciousness. But in general, (3) the properties of a whole are determined by the properties of the particles that compose it, and (4) beings like us are composed of physical particles. Since no combination of physical particles could combine to produce consciousness, it follows that beings like us must not be conscious, yet clearly (5) beings like us are conscious.

The foregoing claims (1)–(5) are jointly inconsistent. Claims (1)–(4) imply that beings like us cannot have conscious experiences, but claim (5) implies that we can. Solving the problem requires rejecting at least one of these claims. Hylomorphists solve the problem by rejecting claim (3). Although properties like the mass, location, and position of a complex whole may be determined by the masses, locations, and positions of the particles that compose it, this is not true of all properties, say hylomorphists. The reason is that some properties depend on structure (Section 6.3). This is enough to dispense with the problem, but there is more to say.

Hylomorphists spot an ambiguity in claim (2). It is true that no number of particles can produce conscious experiences by themselves, but if those particles are suitably structured, if they manifest their powers in coordinated ways, then there is a sense in which a number of nonconscious particles can combine to produce a conscious whole. For on the hylomorphic view, a conscious experience is composed of the coordinated manifestation of the powers of something's parts and surrounding materials. Consequently, there is a sense in which hylomorphists can reject claim (2) as well.

13.3 Hylomorphism and the Problem of Downward Causation

Hylomorphists seem initially to be in the same boat as emergentists when it comes to the problem of downward causation (Jaworski 2011: 344–52; 2012). Recall the version of the problem formulated in Section 13.1:

(1) Actions have mental causes.
(2) Actions have physical causes.
(3) The mental and physical causes of actions are distinct.
(4) If actions have multiple causes, then they are overdetermined.
(5) Actions are not overdetermined.

Because hylomorphists endorse the existence of emergent properties with the characteristics described in Section 6.3, they cannot reject claims (1) or (3), nor can they reject claim (2), for as we've seen, they are committed to the idea that higher-level behavior never violates lower-level physical laws (Section 8.2). That leaves claims (4) and (5). Between them, hylomorphists target (4).

According to hylomorphists, emergentists face a problem with mental causation because they implicitly assume that there is only one kind of causal relation—the kind that can be described by physics. Even nonphysical emergent powers are supposed to operate like emergent forces on the emergentist view.[3] Consequently, if mental states are to influence behavior, they must do so in the way physical forces do. Roger Sperry describes the idea in the following terms:

> most of the atoms on our planet are primarily moved around not by atomic or subatomic laws and forces, as quantum physics would have it, but rather by the laws and forces of classical physics, of biology, geology, meteorology, even sociology, politics, and the like. (1984: 201)

According to Sperry, emergent properties are emergent forces. Consequently, if the mental causes of actions do not exclude physical causes, as the denial of (2) would have it, the mental and the physical causes of actions must together overdetermine their effects. But hylomorphists reject the assumption that there is only one kind of causal relation. We saw in Chapter 10 that according to them, the term 'cause' is equivocal: it can be used to refer to many different kinds of causal relations. Once we clarify how it is used in claims (1)–(5), we see either that those claims are all true but consistent, or else that one of them—claim (4)—is false, but its falsity does not have any awkward consequences.

We saw in Section 10.3 that one way of giving content to the idea that there are different kinds of causal or explanatory relations is to view causes or explanations as answers to certain kinds of questions. This view was endorsed by Aristotle (*Physics* 194b16–20), and more recently by van Fraassen (1980: 134). We distinguished

[3] See McLaughlin (1992) for more on this point.

mechanistic explanations, the sort provided by functional analysis, from explanations of living behavior that appeal to people's reasons. That distinction appears in Plato (*Phaedo* 98c–99b), and also in Davidson (1970), Dretske (1988), Burge (1993; 2006), and others. As an empirical matter of fact, there appear to be at least two different ways of explaining human behavior: a way that appeals to reasons, and a way that appeals to physiological mechanisms. Because hylomorphists understand causes and causal relations by appeal to explanatory factors and explanatory relations, they take these two kinds of explanation to correspond to two kinds of causation. Beliefs and desires cause or contribute to actions in one kind of way—they *rationalize* actions—and neural events cause or contribute to actions in a different kind of way—they *trigger* the muscular subsystems involved in actions. This distinction among kinds of causal relations enables hylomorphists to solve the problem of mental causation in an attractive way.

Because the term 'cause' is used in different ways, say hylomorphists, claims (1), (2), and (3) must be rewritten:

(1′) Actions are rationalized by thoughts, feelings, and/or perceptions.
(2′) Muscular contractions are triggered by events in the nervous system.
(3′) Rationalizing causes and physiological triggers are distinct.

These claims are jointly consistent, and hylomorphists can endorse all three. But hylomorphists reject claim (4). Because reasons and physiological mechanisms contribute to actions in different ways, they cannot be overdetermining causes of actions. Consider an analogy with the car crash example discussed earlier (Section 8.5). Because blood-alcohol levels and shallow roadway gradings contribute to crashes in different ways, they are not overdetermining causes of crashes. Overdetermination in the relevant sense implies that the overdetermining causes contribute to their effects in the same way—each is by itself a sufficient cause. But that is not the case here. According to hylomorphists, human behavior is a highly structured phenomenon that comprises a complex range of causal factors. Explanations that appeal to reasons and explanations that appeal to physiological mechanisms pick out causal factors of different sorts: they answer requests for different kinds of information, just as explanations that appeal to roadway grading and explanations that appeal to blood-alcohol levels provide different kinds of information about car crashes. There is thus no threat of actions being overdetermined by reasons and physiological triggers. Claim (4) is thus false.

Hylomorphists say that in light of the triggering/rationalizing distinction, claim (5) can be true only if it is rewritten. Here are two possible ways of doing so:

(5′) An action does not have more than one rationalizing cause.
(5″) An action does not have more than one physiological trigger.

Hylomorphists are free to endorse both claims, and they are also free to reject them. Either way, they argue, they succeed in solving the problem, for (5′) and (5″) are both

consistent with (1′)–(3′), and so are the denials of (5′) and (5″). Consequently, whether hylomorphists endorse (5′) and (5″) or deny either, they end up solving the problem.

Several remarks are in order about this solution. First, the overall strategy of positing noncompeting causes for actions is not new. Consider Tyler Burge:

> I doubt that there is any rational ground to think that a belief that a physical effect has both a mental cause and a physical cause forces a choice between maintaining that the causes are 'the same' and maintaining that they are 'in competition' or 'in tension' . . . Kim asks, given that the physical cause is 'sufficient', what 'work' remains for the mental cause . . . I am inclined to think that the question trades on unclarified notions of sufficiency and work . . . The idea of physical cause using up all the work so that there is no work left for the mental cause . . . seems to trade on a kind of hydraulic model. According to this model, so much energy is needed to get the job done. Given that enough energy is expended to get the job done by the physical cause, the mental cause is left without any need to expend its energy. This idea . . . seems to import a conception of the relation between physical and mental causation that is not sanctioned by ordinary explanations in the physical and human sciences . . . The notion of work has homes both in physics and in talk of physical labor. The notion may elicit thinking of the psychological cause as like a further physical cause, offering an extra infusion of energy that is in fact not needed to supplement an already sufficient physical cause. Here mental causation would be implicitly regarded as a form of physical causation. In such a role it can easily seem to be an intrusive, competing, physical-like cause. (2006: 379–80)[4]

If there is a difference between the hylomorphic solution and Burge's, it is that hylomorphists situate their solution within an overall metaphysics which makes it clear why there would be different kinds of noncompeting causal factors that contribute to actions. Burge's own basis for endorsing a pluralism of causes is explanatory practice, but Kim has argued that this basis is insufficient to address the problem:

> As Burge says, our confidence in the truth of familiar intentional explanations does exceed our commitment to any recondite metaphysical principles . . . But I doubt that very many of us who have "worried" about mental causation have actually been concerned about the possibility that our thoughts and desires might turn out to have no powers to move our limbs. Our worries are not *evidential* or *epistemological* worries . . . [T]he problem of mental causation is primarily a metaphysical problem. It is the problem of showing *how* mental causation is possible, not *whether* it is possible . . . [T]he how-question of mental causation arises because there are certain other commitments . . . that make mental causation prima facie problematic . . . The issue is *how to make our metaphysics consistent with mental causation*, and the choice that we need to make is between various *metaphysical alternatives*. (1998: 61–2)

Because hylomorphists go beyond Burge in working out a detailed metaphysics, their view is immune to Kim's criticism.

4 John Dupré (1993: 100–1) says something similar.

Second, the hylomorphic solution is a species of what Kim (1991) has called a *dual-explanandum* or *two explananda* strategy. Rationalizing and triggering explanations do not have one and the same thing as a shared explanandum; they rather address two different explananda. Rationalizing explanations address actions; triggering explanations address the physiological mechanisms whose powers, when suitably coordinated, compose actions. This distinction is evident in the rewritten claims (1′) and (2′). The former is a claim about actions; the latter, a claim about muscular contractions. The original formulation of claims (1) and (2) masks this distinction—one implied by a commitment to activity-making structures (Section 8.1). Recall that according to hylomorphists, muscular contractions, neural firings, and other physiological events compose actions only if they are coordinated in the right kinds of ways—only if they have the right activity-making structure. To identify actions with lower-level physiological events, say hylomorphists, is to deny that there are activity-making structures. Since hylomorphists are committed to the existence of activity-making structures, they deny that actions are identical to the physiological occurrences that compose them.

It is worth contrasting the hylomorphic solution with two other dual-explanandum solutions. The first is advanced independently by Norman Malcolm and A. I. Melden:

> [Purposive and neural explanations of behavior] explain different things. Purposive explanations explain actions. Neurophysiological explanations explain movements. Both explain behavior: but we can say this only because we use the latter word ambiguously to cover both actions and movements . . . the two kinds of explanation belong to different "bodies of discourse" or to different "language games." They employ different concepts and assumptions. One kind of explanation relates behavior to causal laws and to concepts of biochemistry and physiology, to nerve pulses and chemical reactions. The other kind of explanation relates behavior to the desires, intentions, goals, and reasons of persons. The two forms of explanation can co-exist, because they are irrelevant to one another. (Malcolm 1968: 52)

> Where we are concerned with causal explanations, with events of which the happenings in question are effects in accordance with some law of causality, to that extent we are not concerned with human actions at all but, at best, with bodily movements or happenings; and where we are concerned with explanations of human action, there causal factors and causal laws in the sense in which, for instance, these terms are employed in the biological sciences are wholly irrelevant to the understanding we seek. The reason is simple, namely, the radically different logical characteristics of the two bodies of discourse we employ in these distinct cases—the different concepts which are applicable to these different orders of inquiry. (Melden 1961: 184)[5]

[5] W. D. Hamlyn draws a similar distinction: "With movements we are concerned with physical phenomena, the laws concerning which are in principle derivable from the laws of physics. But the behaviour which we call 'posting a letter' or 'kicking a ball' involves a very complex series of movements, and the same movements will not be exhibited on all occasions on which we should describe the behavior in the same way. No fixed criteria can be laid down which will enable us to decide what series of movements shall constitute 'posting a letter.' Rather we have learnt to interpret a varying range of movements as

The hylomorphic approach to mental causation is similar to Malcolm's and Melden's in one respect: it posits two different kinds of explanations and two different explananda. It nevertheless rejects their tacit commitment to causal monism. Malcolm and Melden implicitly assume that causal explanation is exclusively scientific. Since they deny that psychological explanations are scientific, they conclude that psychological explanations are not causal. Hylomorphists disagree. They side instead with Davidson (1963) that psychological explanations—rationalizations in particular—are causal explanations, although they do not accept Davidson's (1970: 208) further assumption that causal relations must be subsumed under strict laws.

A second dual-explanandum strategy is endorsed by Fred Dretske (1991), who recommends Jaegwon Kim's description of his view:

> The general approach [Dretske] adopts can be called "the dual-explanandum strategy": Resolve the explanatory rivalry [between rationalizations and physiological explanations] by holding that the two explanations do not in reality share the same explanandum ... [R]ationalizations explain *behavior* ... not bodily movements ... Dretske tells us ... to identify behavior not with motor output M ... but rather with the relational structure, *C's causing M* ... [B]ehaviors are *causings*, and explanations of behavior must explain these causings by providing *causes of causings*. Neurobiology has M, the "product" or "result" of behavior, as its proper explanandum; it can also explain *how*, that is, through what sequence of intervening states, C led to the production of M. However, it is the proper job of rationalizations to explain *why C caused M*. (Kim 1991: 293–4)

Hylomorphists agree with Dretske that the explananda of rationalizations are not mere bodily movements, and they agree that bodily movements are triggered by internal states of some sort. They are nevertheless not committed to claiming, as Dretske is, that the job of rationalizations is to explain why or how the causal connections between the internal states and bodily movements got established. The door-locking example in Section 8.4 illustrates that rationalization likely involves a different kind of pattern discernment. Understanding why Gabriel is locking the door does not seem to involve grasping the natural selective or environmental factors that yoked his neural mechanisms to his bodily movements.

Another thing to note about the hylomorphic solution: it implies that the arrow diagrams made popular by Jaegwon Kim for discussing the problem of mental causation are misleading. Figure 13.2 includes three examples of Kim-style mental causation diagrams. The arrows represent causal relations, and the M, P, and A represent a mental event, a physical event, and an action, respectively. Kim's diagrams have been enormously influential, but according to hylomorphists, they misrepresent human action in at least three ways. First, they leave out the social and environmental factors that are sometimes necessary for action to occur (Section 8.1).

coming up to the rough standard which we observe in acknowledging a correct description of such behaviour as posting a letter" (Hamlyn 1953: 134–5).

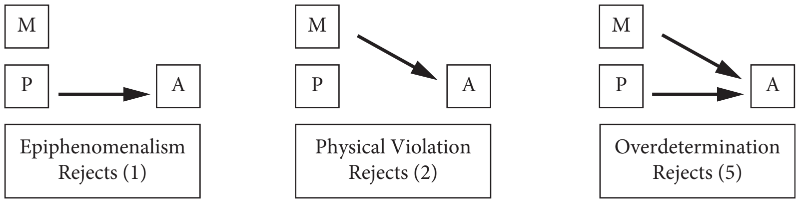
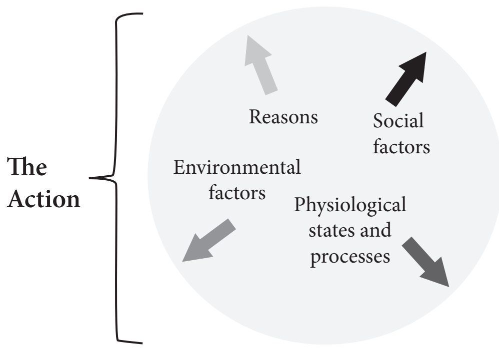


Figure 13.2 Kim-style mental causation diagrams.

Figure 13.3 A new depiction of mental causation.



Second, they suggest that there is only one kind of causal relation which is represented by the arrows in the diagram. Finally, the arrows suggest that the kind of causal relation involved is a triggering relation. They do not capture the distinctive character of rationalizations. Kim-style diagrams might be well suited to representing physiological events that trigger other physiological events, say hylomorphists, but they misrepresent the nature of actions. A better representation might be Figure 13.3, which depicts physiological conditions, along with social, psychological, and environmental ones, as necessary factors that contribute to the occurrence of an action in different ways.

13.4 Objections to the Hylomorphic Solution

Consider now an objection to the hylomorphic solution to the problem of downward causation (Jaworski 2011: 348–52; 2012: 164–8). It claims that hylomorphists provide no solution at all, for they are forced either to deny that rationalizing causes exercise any real control over human behavior, or else to concede that rationalizing causes and physiological triggers overdetermine actions. Why? The reason, says the objection, is that a thing's properties are metaphysically necessitated or determined by the properties of its parts. Your properties, for instance, including your psychological properties, are determined by the properties of your organic parts, such as your brain. Furthermore, the properties of your brain are determined by the properties of the

molecules that compose it, and their properties are determined by the properties of the atoms that compose them, and so on. Ultimately, then, the properties of everything are determined by the properties of fundamental physical particles since these are the basic constituents of everything. Now if everything is determined by what happens at a fundamental physical level, then beliefs, desires, and other mental states do not really contribute to explanations of what people do; the real explanations for human behavior and everything else are given by fundamental physics. Beliefs, desires, and other mental states might rationalize actions, as hylomorphists claim, but these rationalizations cannot have any real causal import, for they do not have explanatory import, and on a hylomorphic view, causes are explanatory factors. Hylomorphists could try to respond by attributing real causal import to rationalizations, but if they do so, they commit themselves to the idea that actions are overdetermined by the rationalizations and the lower-level determining factors. Consequently, says the objection, hylomorphism does no better vis-à-vis the problem of mental causation than competing theories like emergentism.

Hylomorphists respond that this objection is based on two assumptions they reject, namely that there are *not* many different kinds of causal relations, and that structure is *not* a basic explanatory principle. As a result, they say, the objection begs the question against them. It does not prove that their solution fails; it rather assumes it from the outset. Let us consider this response in detail.

The objection's first tendentious assumption, say hylomorphists, is that higher-level and lower-level causes—reasons and triggers, in this case—compete to occupy a single causal role. Exponents of the objection assume that higher-level and lower-level causes would have to explain their effects in the same ways, and consequently, they assume that higher- and lower-level causes would either have to exclude each other from occupying a single causal role (the role of being the one and only cause of higher-level behavior), or else they would have to share that role and be redundant overdetermining causes of higher-level behavior. We have seen, however, that hylomorphists reject this understanding of causes and causal relations in favor of a plurality of causes and causal relations. There is not a single causal role that rationalizing causes and triggers compete to occupy, they say. Actions are complex multistructure phenomena: each comprises many levels of structural complexity and many different causal factors. Consequently, explaining an action completely would require a description of a broad range of social and environmental factors, as well as subactivities and substructures at every level, with causes at each level contributing to an explanation of the action in different ways. Typically, however, we are not interested in explaining actions completely. We instead focus on some kinds of explanatory factors to the exclusion of others. In pedestrian contexts, for instance, we are typically interested in the rational structure of people's behavior: we focus on their reasons for doing this or that. In neuroscientific contexts, by contrast, we are not interested in people's reasons, but in the neural substructures that enable rational behavior to occur. Because causes of the foregoing sorts—rational and

neuroscientific—contribute to people's behavior in different ways, they do not compete to occupy a single explanatory role; they instead play different noncompeting roles in explaining a complex phenomenon. The objection assumes at the outset that this kind of causal pluralism is false, that human intentional actions are not complex multistructure phenomena that comprise many different kinds of causes and causal relations. As a result, say hylomorphists, the objection begs the question against them. Moreover, since hylomorphists take their causal pluralism to be grounded in our best empirical explanations of why and how things operate as they do, they urge their opponents to show that the empirical data do not support the pluralistic view.

The objection's second tendentious assumption, say hylomorphists, is that all of a thing's properties are metaphysically determined by the properties of its parts. Earlier, we called this the *lower-level determination thesis* (Section 10.2):

> *Lower-level determination*: Necessarily, for any structured individual, x, if x engages in activity A at time t, then (a) x has proper parts and surrounding materials with powers, $P_1, P_2, \ldots, P_m$, such that the manifestations of $P_1, P_2, \ldots, P_m$ at t explain x's A-ing at t, and (b) necessarily, for any individual z and time t^*, if z has proper parts and surrounding materials at t^* exactly similar to x's at t, with powers, $Q_1, Q_2, \ldots, Q_m$, exactly similar to the powers of x's proper parts and surrounding materials, then z engages in A-ing at t^*, and the manifestations of $Q_1, Q_2, \ldots, Q_m$ at t^* explain z's A-ing at t^*.

Determination is a type of necessitation relation; it conjoins necessitation with explanation: if F-things determine G-things, then something's F-properties necessitate its G-properties, and its F-properties also explain its G-properties. According to exponents of lower-level determination, if Alexander is composed of lower-level components with lower-level properties $F_1, F_2, \ldots, F_n$, then Alexander is guaranteed to have certain higher-level properties precisely because his lower-level components have $F_1, F_2, \ldots, F_n$. Because all of Alexander's properties are explained by lower-level conditions, exponents of this view think that it undermines the explanatory status of higher-level properties.

We saw in Chapter 10 that hylomorphists reject lower-level determination. The reason is that lower-level determination is incompatible with the claim that structure is a basic explanatory principle. According to hylomorphists, it might be plausible to endorse something like lower-level determination when it comes to what William Wimsatt (1985) calls *aggregative properties* such as mass. It is plausible to suppose that, say, an organism's mass is determined by the masses of its fundamental physical constituents. But not all properties are like this on the hylomorphic view. Some depend not just on the materials that compose a thing, say hylomorphists, but on the way those materials are organized or structured. To deny this in favor of lower-level determination for all properties is to deny at the outset that hylomorphism is true. The objection thus begs the question against hylomorphism in a second way.

In response, exponents of lower-level determination could say two things. First, they could expand the lower-level determination thesis to include the lower-level conditions that constitute an individual's environment. Second, in defense of lower-level determination theses in general, they could argue that the best explanation for the supervenience of higher-level properties on lower-level properties is that lower-level properties determine higher-level ones. If lower-level conditions determine higher-level conditions, then indistinguishable lower-level conditions will yield indistinguishable higher-level conditions. Lower-level twins will be guaranteed to be higher-level twins. Hence, exponents of lower-level determination could say, the supervenience of higher-level conditions on a suitably broad base of lower-level conditions provides us with good reason to believe that some version of the lower-level determination thesis is true.

Hylomorphists can argue in response that their view implies strong supervenience without implying lower-level determination, and as a result, lower-level determination does not provide an obviously superior explanation for supervenience relations. We've seen that according to hylomorphists, higher-level conditions supervene on lower-level physical conditions, but these supervenience relations are underwritten not by lower-level determination, but instead by material dependence: higher-level structural phenomena require that specific lower-level conditions be satisfied (Section 10.4). To understand how this works, imagine two scenarios. In Scenario A, Alexander believes that there are eight planets in our solar system, and in Scenario B, he believes that there are nine. According to hylomorphists, this difference in Alexander's beliefs likely depends on social or environmental differences in the two circumstances. It is plausible to suppose, for instance, that differences in beliefs about the planets depend on receiving a certain kind of education, and so Alexander's belief in Scenario A must be due to something in his education (a lecture on astronomy, say) that is missing from Scenario B. Hylomorphists are free to claim that this social or environmental difference depends in turn on lower-level physical differences. If the lecture occurs in Scenario A but not in Scenario B, then in one scenario but not the other, the lecturer's utterances will vibrate the air molecules that eventually impinge on Alexander's eardrums. Without some sort of lower-level physical differences between the two scenarios, there could not be the social or environmental differences that are necessary for the differences in Alexander's beliefs.

According to hylomorphists, then, higher-level conditions are not determined by lower-level conditions, they merely depend on lower-level conditions: different higher-level conditions depend on different lower-level ones. Having a belief that there are eight planets requires lower-level conditions different from having a belief that there are nine. Consequently, Alexander's beliefs about the number of planets cannot differ in Scenarios A and B unless those scenarios differ physically. It is by no means obvious, say hylomorphists, that this is a weaker explanation for supervenience than lower-level determination. At the very least, exponents of lower-level

determination must show that their explanation is superior in order to support this part of their objection.

According to hylomorphists, therefore, the objection to their view is based on assumptions about causation and structure that they reject. It represents not an independently warranted argument against their view, so much as an alternative approach to downward causation. That approach, they say, is responsible for generating the problem of downward causation in the first place—something amply illustrated by the line of reasoning the objection follows. Part of what makes their own approach so attractive by contrast, hylomorphists insist, is that it avoids generating problems of this sort.

A similar objection to the hylomorphic solution is advanced by Howard Robinson:

> It seems that modern hylomorphism faces the following dilemma . . .
>
> (1) The concept of structure is essential to all or most natural sciences.
> (2) If something is essential to a valid mode of explanation or understanding, then it should be conceived realistically. So,
> (3) There are real structures in the world.
>
> But the following also has appeal:
>
> (4) [a] It is sufficient for the concept of structure to be applicable that elements be appropriately related in the world, and [b] these relations can be characterized without using the notion of structure . . . by specifying the spatio-temporal location of the elements and their causal influence on each other.
>
> If (4) is correct, it looks as if . . . structures are not part of the basic furniture of the world . . . [T]hey contribute nothing over and above the 'forces' of physics, and . . . are nothing above their constituents and their spatiotemporal and causal relations. (2014: 209–10)

What I've labeled clause [a] of Premise (4) is ambiguous, but on one plausible interpretation it is implied by the structo-physical necessitation thesis defended in Section 9.3. If that thesis is true, then physical twins must be structural twins. There can be no structural differences between x and y without there being corresponding physical differences between them, and in that sense, a complete physical description of x and y can be sufficient for knowing how matters stand with them structurally: if x and y are exactly similar physically, they must be exactly similar structurally. Moreover, hylomorphists needn't balk at clause [b]. They are free to claim that there are spatiotemporal and causal relations that can be described exhaustively by physics. What they balk at is the suggestion that a description of spatiotemporal and causal relations at the level of physics is sufficient for providing a description of higher-level structure.

The claim that the conceptual resources of physics are adequate to the task of describing higher-level structure is the basis of Williams' Worry 1, the worry that higher-level structures are just complex relations among physical particles which can be exhaustively described and explained by physics. We considered the hylomorphic response to this worry in Section 12.2. Hylomorphists endorse structure realism; as a result, they deny that higher-level biological, psychological, and other structures are

the kinds of things that physics can exhaustively describe. It is true that those structures supervene on properties or relations that can be exhaustively described by physics, but according to hylomorphists, none of those properties or relations is the kind of dynamic higher-level structure that properly constitutes the subject-matter of biology, psychology, or other special sciences. Hylomorphists could be wrong about this: perhaps in the long run it will be possible to identify biological, psychological, and other higher-level structures with things that can be described exhaustively by physics, and in that case, hylomorphism will in fact turn out to be false. But Robinson provides no reason to think the sought-after identifications are forthcoming.

Hylomorphists will also balk at Robinson's final premise, which he doesn't number, the claim that if (4) is correct, then structures contribute nothing over and above what physics describes, and are nothing above the physical constituents of things and their physical relations. This is a claim that hylomorphists reject on account of their commitment to structure realism (Section 1.2). Is there any reason to think that hylomorphists are wrong and that Robinson's unnumbered premise is right? In support of the premise, he suggests the following:

> [Jaworski] is a *causal pluralist*, but [his position] may seem self-undermining . . . [His] avowal of causal pluralism is immediately followed by the assertion that "all *forces* are operating at a fundamental physical level" . . . What one has . . . is an explanatory pluralism, with causation adopted into the domain of explanation; the wholly external mind-free element is *force* and this is exclusively micro. So Jaworski is really only claiming that, once one realizes that most explanations are causal explanations, explanatory pluralism is pluralism enough to constitute or ground a full causal realism about all levels . . . [O]ne may wonder whether *explanatory* pluralism is enough to justify or constitute the real efficacy of the non-basic levels. The importance of *explanations* at the level of complex structures is not disputed, but whether these represent just ways of *conceptualizing* a fundamental level that we cannot normally access is another matter. (2014: 208)

According to Robinson, the kind of hylomorphic theory I defend is guilty of trading in a pluralism of objective causal factors for a pluralism of subjective explanatory modes. By his lights, someone might concede that there are many different kinds of explanations, but claim that these explanations are just different ways of conceptualizing phenomena that can be described and explained exhaustively at the level of fundamental physics.

If Robinson is right, then the hylomorphic view of explanation is really no different from the nonreductive physicalist view of explanation described in Section 11.4. On that view, explanations are simply responses to subjective interests that underwrite intuitive judgments about what kinds of descriptions and explanations are satisfactory, and the kinds of descriptions and explanations we find intuitively satisfying needn't correspond to the objective factors that are responsible for producing real-world effects. What motivates this view of explanation is the nonreductive physicalist

commitment to the exhaustiveness of physical explanation. If physics is capable in principle of providing an exhaustive account of everything, and the explanatory categories of the special sciences do not correspond to those of physics, as nonreductivists claim, then the only way of countenancing the legitimacy of special scientific explanations is to say that their legitimacy derives from a source different from that of explanations in physics. Because physicalists in general insist that the latter explanations are legitimate by virtue of describing the objective factors that are responsible for things, nonreductivists are committed to grounding the legitimacy of special scientific explanations in something else, such as our intuitive judgments of what counts as satisfying.

We saw in Section 11.4 that hylomorphists reject this view of explanation. According to them, explanations pick out objective causal factors. A subjective element enters the process of explanation because the natural world is replete with such factors, and consequently, our concrete acts of explaining must select some of those factors for attention while bracketing others. Which factors we select is doubtless a function of our subjective interests. But making a selection among objective causal factors does not make them or their contributions to the effect any less objective: it does not turn them into subjective, internal, or mind-dependent occurrences. The balding tires, the faulty break mechanism, the shallow roadway grading, the high blood-alcohol level—these and other factors contribute to the car crash in objective, external, mind-independent ways. Choosing to focus on one factor or the other doesn't change that.

Nor does this picture commit hylomorphists to claiming that explanations are "Platonic entities, like propositions," as Robinson claims (2014: 213). The explanatory factors themselves are not Platonic entities. The car's crashing, the brakes' being faulty, the driver's high blood-alcohol level, and the other contributing factors are all events or tropes (Chapter 3): individuals having properties or standing in relations at times. The same is true of concrete acts of explanation. These are acts in which individuals engage—acts which are, on the hylomorphic view, essentially embodied in the parts that those individuals possess (Sections 8.2–8.3). It is incorrect to claim of either the explanatory factors or the explanatory acts themselves that they are "not part of the world they explain," but merely, "part of the . . . conceptual structure we apply" (213). If the world consists of individuals having particularized properties, then the explanation that the car crashed because its brakes were faulty picks out mind-independent constituents of the world and a mind-independent causal relation between them.

In light of the foregoing considerations, it is difficult to appreciate Robinson's claim that the hylomorphic view is "self-undermining." Forces are causes of only one sort on that view. There are causes and causal relations of other sorts, including ones described by the special sciences. Robinson takes this suggestion to imply that these so-called other causes are not really causes at all, but mere explanatory postulates—"parts of explanatory schemes," as he puts it. What he fails to explain is why we should believe this claim. Prima facie it seems implausible. The idea that the

explanatory schemes of the special sciences do not express genuine causes but that the explanatory schemes of physics do, introduces an asymmetry in the general picture of how explanations are related to causes. If explanations in general map onto causes, and different kinds of explanations map onto different kinds of causes, then there is a unified account of the relation between causes and explanations. Robinson suggests something different: only some kinds of explanatory schemes map onto causes. But why? What justifies this claim? I can detect nothing in Robinson's remarks that answers this question. An answer would follow straightaway if the only causes that existed were forces—if, in other words, the causal pluralism that hylomorphists endorse were false. But then why should we accept this claim? Again, I can detect nothing in Robinson's remarks that answers this question. As a result, I can detect nothing in Robinson's remarks that would undermine the claim that structure realism is compatible with all forces operating at the level of physics.

Robinson's argument appears to be flawed in at least two ways therefore. First, to the extent that it attributes to hylomorphists a view of explanation they do not endorse, the argument attacks a straw man. Second, to the extent that it takes for granted a view of explanation that hylomorphists reject, the argument begs the question. I conclude that Robinson's argument does not carry the day.

13.5 Worries 6 and 7: Epiphenomenalism

The hylomorphic approach to downward causation has implications for the sixth and seventh versions of Williams' worry. Worry 6 says that hylomorphism must be committed to some type of epiphenomenalism. According to hylomorphists, the behavior of structured individuals like us never violates the laws governing their fundamental physical constituents (Section 8.2). But if structured individuals like us can never violate fundamental physical laws, then it is impossible for our thoughts and feelings to make any causal difference to our behavior. How, after all, could those thoughts and feelings make a difference if not by intervening at the level of fundamental physics? Hylomorphists must therefore be committed to something like epiphenomenalism. Even if there are structures that elude physical description, our behavior can ultimately be explained exhaustively by physics alone.

Hylomorphists respond that Worry 6 is based on an assumption they reject. It tacitly assumes that in order to make a difference to things, higher-level causal relations must conform to the models of causation described by physics. But according to hylomorphists, there is not just one kind of causal relation, there are many kinds. Consequently, there is no reason to think that thoughts and feelings are epiphenomenal simply because they do not conform to the models of causation operative in physics.

Importantly, this kind of causal pluralism does not imply that higher-level phenomena are able to make a difference in the world without making a difference at the level of fundamental physics. Every higher-level difference must correspond to a

lower-level difference. This is implied by hylomorphists' commitment to strong supervenience: necessarily, there can be no higher-level differences between worlds w_1 and w_2 without there being lower-level differences between w_1 and w_2. These correlations between higher- and lower-level differences, however, do not depend on higher-level phenomena conforming to models of causation in physics.

Worry 7 pushes things a step further. It assumes that Laplacean determination is true:

> *Laplacean determination*: Necessarily, for any total state, *S*, of the universe at time t_1, there is exactly one possible resultant state, S^*, that comes about at time t_2 because of *S* in conjunction with the laws of nature.

Laplacean determination is compatible with hylomorphism, so the assumption is unobjectionable on hylomorphic grounds. If Laplacean determination is true, then it seems that whatever states of affairs come about in the universe come about because of laws of nature and antecedent conditions, and if that is true, then it seems that higher-level structural phenomena have no explanatory role to play in the world: they are mere epiphenomena.

Hylomorphists respond that the conclusion of Worry 7 does not follow from the premises. Suppose that I throw a baseball at time *t*. Since throwing a baseball is a structured activity, it has an activity-making structure essentially. My throwing a baseball at *t* would not exist were it not for the activity-making structure that unifies the manifestations of diverse powers into a single activity. Suppose now that Laplacean determination is true and that the probability of my throwing a baseball at *t*, given antecedent conditions and the laws of nature, is 1; given those conditions, it was not possible for me not to have thrown a baseball at *t*. It does not follow from this that the throwing-a-baseball-wise structure of my activity is not essential to it. Nor does it follow that the activity does not produce effects of its own, such as striking the first baseman's glove or making an out. All that follows if Laplacean determination is true is that these effects are produced with a probability of 1. Nothing follows about the inefficacy of activities or activity-making structures.

Someone might look to reassert the worry by focusing on lower-level occurrences. If Laplacean determination is true, then it seems that lower-level occurrences must bring about other lower-level occurrences with a probability of 1. If hylomorphism is true, moreover, then lower-level occurrences necessitate higher-level occurrences (Chapter 9). Consequently, it seems as though higher-level occurrences must come about because of lower-level occurrences and not because of higher-level occurrences.

Once again, hylomorphists respond that the conclusion of the argument does not follow. In order to derive the conclusion that higher-level occurrences must come about because of lower-level occurrences and not because of higher-level ones, the argument must assume something like the lower-level determination thesis (Sections 10.2 and 13.4): it must assume that lower-level occurrences do not merely necessitate higher-level occurrences, but that lower-level occurrences also explain them. In fact,

the argument must assume more than that: it must assume that if higher-level occurrences are explained by lower-level ones, then those higher-level occurrences cannot be explained by anything else; it must assume that there is a single explanatory role to be filled, and that if lower-level occurrences fill that role, higher-level ones cannot. We have seen, however, that hylomorphists reject both assumptions (Section 13.4). They deny lower-level determination, and they deny that there is only one kind of causal factor or causal relation. On the one hand, say hylomorphists, there are structures, basic ontological and explanatory principles, that factor essentially into what things are and what they can do. On the other hand, there are lower-level materials. These also factor essentially into what things are and what they can do, but they do so in a different way—the kind of way revealed through functional analysis and expressed in mechanistic explanations of higher-level behavior. The difference between this latter kind of explanatory contribution and the kind made by structure or organization is the reason, recall, that higher- and lower-level occurrences do not overdetermine what they explain. It is also the reason higher- and lower-level occurrences do not compete to occupy a single explanatory role. The assumption that they do compete to occupy a single explanatory role is, however, the basis of the worry. Consequently, if hylomorphism is true, the worry cannot arise.

13.6 Hylomorphism and the Problem of Other Minds

Hylomorphism also has implications for another mind-body problem: the problem of other minds (Jaworski 2011: 331–8). The problem of other minds arises on account of two crucial assumptions. The first is that we cannot directly perceive other people's thoughts, feelings, and perceptions; that our knowledge of them depends on making inferences from bodily behavior. The second assumption is that the bodily behavior we directly observe can be correlated with many different kinds of thoughts, feelings, and perceptions. A grimace could be due to a genuine feeling of pain, or to a desire to deceive you into believing that I am in pain. Because my movements and gestures do not provide decisive evidence that I have these or those thoughts, feelings, and perceptions, you have difficulty knowing what my thoughts, feelings, or perceptions are. In fact, you might have difficulty knowing whether I have any thoughts, feelings, and perceptions at all. If this is in fact possible, then you cannot infer on the basis of my bodily movements and gestures what thoughts, feelings, and perceptions I have or whether I have any thoughts, feelings, or perceptions at all. Consequently, if both of the foregoing assumptions are true, there is a real problem knowing other people's thoughts, feelings, and perceptions. According to hylomorphists, however, both assumptions are false.

The hylomorphic approach to other minds depends on three claims:

(1) Thoughts, feelings, and perceptions are coordinated manifestations of the powers of our parts and surrounding materials.

(2) We can perceive some of these coordinated manifestations.
(3) Among the coordinated manifestations we can perceive, some fix the referents of the predicates and terms we use to describe thoughts, feelings, and perceptions.

Given reasonable assumptions, these claims imply that we can know other people's thoughts, feelings, and perceptions in something analogous to the way we know that things in the environment are instances of certain natural kinds (Kripke 1972; Putnam 1975c). We learn to use natural kind terms by applying them to paradigmatic instances of the kind. We apply the term 'water' to the stuff in this bottle, to the stuff that fills this lake, to the stuff that falls in droplets from that cloud, and so on. These cases fix the referent of 'water': they determine which stuff in the environment 'water' refers to. Some of these referent-fixing features constitute our initial conception of water—our initial understanding of what water is: a colorless, odorless, thirst-quenching liquid that fills lakes and rains from clouds. Later on, when we study water scientifically, we might discover other properties it has, but initially, our conception of water is determined by our pedestrian interactions with paradigmatic instances of it.

According to hylomorphists, something analogous is true of the predicates and terms we use to describe thoughts, feelings, and perceptions. We see things, feel them, recognize and want them, strive for them, and because we are social animals, we have evolved abilities to make our seeings, feelings, recognizings, and wantings known to others. Sometimes, we make them known by movements, facial expressions, and vocal cues—"body language" in general. Other times, we communicate them by means of symbols in a written or spoken language. These symbolic modes of expression are extensions of our natural emotional, perceptual, cognitive, and conative capacities. They are forms of symbol-using behaviors that are grafted onto other forms of expressive behavior. What these symbols express are coordinated manifestations of our powers and those of surrounding materials—patterns of social and environmental interaction. We learn what anger, joy, and fear are through pedestrian interactions in which we are angry, joyful, or afraid. These interactions furnish conditions for correctly applying psychological predicates, and they fix the referents of psychological terms. An example can help illustrate this idea.

Imagine a mother and her small child walking near some shrubbery when a cat unexpectedly jumps out from behind a bush and startles them. They jump back, their eyes widened, brows raised, lips stretched back toward their ears, their hearts race, and the child cries. "Did that scare you?" the mother says, as she picks up the child to comfort her. This episode provides a paradigmatic instance of fear. The child and her mother are both attending to a limited set of perceived social, environmental, and physiological factors: the startling occurrence, the physiological responses to it, and social responses such as the facial expressions, the crying, and the mother's comforting reaction. These factors constitute a set of conditions for correctly deploying the predicate 'scared.' Later, the child imitates the utterance and the intention behind it in

contexts which resemble the paradigmatic case, and comes gradually to master the predicate's use. Later still, the child learns to use terms like 'fear' and predicates like 'is afraid' to describe the condition of being scared. The latter predicates and terms are symbols whose usage builds on previously mastered verbal skills in something analogous to the way our ability to solve quadratic equations builds on previously mastered arithmetic skills. Those previously mastered verbal skills are themselves grounded in forms of social and environmental interaction in which animals like us are prelinguistically competent to engage—coordinated manifestations of powers we humans possess, such as joint attention and intention-reading, which are described here by the psychologist Michael Tomasello:

> [Word learning] emerges naturally from situations in which children are engaged in social interactions in which they are attempting to understand and interpret adult communicative intentions as expressed in utterances . . . [There are] two . . . aspects of the word learning process: (1) the structured social world into which children are born—full of scripts, routines, social games, and other patterned cultural interactions; and (2) children's . . . capacities for tuning into and participating in this structured social world—especially [capacities for] joint attention and intention-reading . . . [H]uman children are born into worlds in which their caregivers have certain activities to perform on a regular basis . . . Some of these routines are fairly constant across cultures (such as nursing) . . . Social interactional routines such as feeding . . . and a host of other activities constitute the formats—joint attentional frames—within which children acquire their earliest linguistic symbols . . . Learning the communicative significance of an individual word consists in the child first discerning the adult's overall communicative intention . . . and then identifying the specific functional role this word is playing in the communicative intention as a whole . . . [T]he shared intentional situation . . . constrains the interpretation of the speaker's communicative intention . . . [C]hildren [thus] acquire linguistic symbols as a kind of by-product of social interaction with adults, in much the same way they learn many other cultural conventions. (2003: 87–90)

What terms like 'anger,' 'joy,' and 'fear' refer to on this account are structured activities, coordinated manifestations of the powers that people and various other things have. According to hylomorphists, this is what thoughts, feelings, and perceptions are: they are phenomena of the same sort as throwing a baseball, coordinated manifestations of the powers of our parts and surrounding materials (Sections 8.3–8.4). Some of these manifestations are perceivable in pedestrian circumstances, others are not. Just as water has properties that we cannot discern in our pedestrian dealings and that do not enter into our initial prescientific conception of water, the same is true of psychological activities: they involve a broad range of conditions, only some of which are apparent to us in our pedestrian interactions. We can directly perceive the facial expressions, gestures, vocal cues, environmental triggers, and social responses, but we cannot directly perceive the movements of ions through cells in the limbic lobe that occur when someone is angry, joyful, or afraid.[6] Conditions of the former, perceptible sort

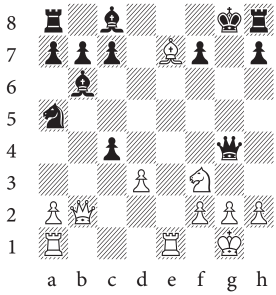
[6] At least the activity of those cells is not grasped by deploying concepts like *LIMBIC LOBE*. Hylomorphism is nevertheless compatible with the claim that in our pedestrian dealings we grasp the activation of cells in the limbic lobe by deploying other, first-personal or pedestrian concepts.

ground our initial abilities to use predicates and terms like 'scared,' 'angry,' 'joy,' and 'fear': they fix their referents and set the conditions for their correct deployment; conditions of the latter, imperceptible sort are discovered through techniques like functional analysis.

This understanding of psychological language has several important implications. First, it implies the falsity of the widely endorsed view that ordinary psychological discourse is like a theory, or that psychological explanation is a species of theoretical explanation. Let us call this view the *theory model* of psychological language. According to the theory model, psychological descriptions and explanations posit unobserved hypothetical entities (thoughts, feelings, and perceptions) whose relations to each other are supposed to explain observed movements, gestures, and other bodily behaviors. When, for instance, we explain why Caesar crossed the Rubicon by saying he *wanted* to secure political power and *believed* marching on Rome the best means of securing it, the theory model claims that we posit hypothetical entities—a belief and a desire—that are related to each other in ways expressed by generalizations like the following: When x wants y, and believes doing z the best means of attaining y, then if nothing inhibits x's pursuit of y, x will generally do z. These posited entities and relations are then supposed to explain why Caesar acted as he did because they subsume Caesar's behavior under the general principle expressed by the psychological generalization. Hylomorphism rejects this understanding of psychological language.[7] Psychological predicates and terms do not refer to or express unobserved hypothetical entities that are posited to explain directly observable human behavior, say hylomorphists; psychological predicates and terms instead refer to or express directly observable manifestations of the powers of individuals and surrounding materials—what I have sometimes referred to as 'patterns of social and environmental interaction' (Jaworski 2011). The way we perceive these patterns is analogous, I would suggest, to the way we perceive patterns in chess or music.

Consider the checkmating pattern in Figure 13.4. The pattern is not hidden from view, it lies in plain sight: anyone who has learned to discern checkmating combinations can see it. Likewise, listeners do not infer that a piece of music has a melody, they perceive the melody directly. Chess playing and music appreciation are possible

[7] The theory model of psychological language was suggested independently by several authors, including Hilary Putnam (1963), Wilfrid Sellars (1963a), Herbert Feigl (1958), and Jerry Fodor (1968a). The principal defenders of the theory model, however, have been eliminative physicalists such as Paul Churchland (1981). Importantly, the theory model of psychological language is not the same as what psychologists and cognitive scientists call 'theory of mind.' The latter term refers to the ability to understand other people's thoughts and feelings. The theory model of psychological language and scientific research on theory of mind have a complex history together, and many psychologists who work on theory of mind implicitly endorse the theory model of psychological language. But the theory model and theory of mind are not the same thing. The ability to understand each other's thoughts and feelings is compatible with rejecting the theory model.



White	Black
1. Q x h8+	K x h8
2. Bf6+	Qg7
3. Re8++	

1. The white queen at b2 takes the rook at h8–Check! The black king takes the queen.

2. The white bishop at e7 moves to f6–Check! The black queen moves to g7.

3. The rook at e1 moves to e8 – Checkmate!

Figure 13.4 Patterns on a chessboard.

because people have the ability to discern patterns on a chessboard, and patterns in sounds. According to hylomorphists, something analogous is true of thoughts, feelings, and perceptions. Normal humans are equipped with the ability to recognize these patterns in each other's behavior. Simple examples of our pattern-discerning abilities include our ability to discern other people's emotional states through facial and vocal expressions.

Paul Ekman (2007) and his colleagues have shown that normal humans have the ability to discern in each other's faces and voices emotional states such as anger, fear, enjoyment, sadness, disgust, and surprise, and that this ability transcends cultural boundaries—facial and vocal recognition of emotional states is a universal human trait. Likewise, Michael Tomasello (2008) and his colleagues have shown that human infants develop the ability to discern other people's intentions as early as nine to twelve months of age.[8] This type of pattern-recognition is not unique to humans, moreover. Other studies show that great apes such as chimpanzees have the ability to discern other people's intentions and perceptions as well.[9]

According to hylomorphists, empirical work like this supports the idea that normal humans are equipped to discern emotions, intentions, and other psychological patterns in each other's behavior. If that is true, then the first assumption motivating the problem of other minds is false. We know other people's thoughts, feelings, and perceptions not by making inferences from bodily movements and gestures to inner hypothetical states, but by directly perceiving the manifestations that fix the referents of psychological predicates and terms. I directly see that you are amused or hear that you are afraid.

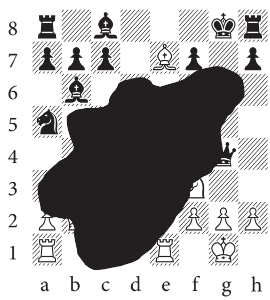
Importantly, the hylomorphic view does not imply that our knowledge of other people's thoughts, feelings, and perceptions is infallible. It looks to accommodate the kind of fallibility we experience in our ordinary dealings with one another while at the same time rejecting the more radical fallibility that motivates the problem of other minds. When it comes to the former point, hylomorphists claim that our

[8] See also Meltzoff 1995 and Tomasello 2003: 26–7.
[9] See Tomasello 2008: 44–7 for a summary of some of this research.

psychological pattern-discerning abilities are subject to error in just the way the exercise of a chess player's pattern-discerning abilities is. A chess player might fail to see mating opportunities in particular cases, and so too, say hylomorphists, we might fail to recognize what people think, feel, or perceive on particular occasions. One reason for this fallibility is that the structured activities we express with psychological predicates and terms involve more than the social, environmental, and physiological factors we initially use to define them. Emotions, for instance, involve more than just movements, facial expressions, and vocal cues; they also involve the activities of physiological substructures which can only be discerned using empirical methods like functional analysis. The activities of these substructures do not factor into our initial prescientific conceptions of anger, joy, or fear, but they are still among the factors whose coordinated manifestations compose those activities. As a result, we can misidentify people's thoughts and feelings in just the way we can misidentify water or gold. We identify water and gold in our pedestrian dealings by appeal to a narrow profile of their properties. As a result, we can mistakenly judge that XYZ is water or that iron pyrite is gold because XYZ and iron pyrite have properties that resemble the properties we typically use to identify water and gold. Something analogous is true of thoughts, feelings, and perceptions on the hylomorphic view. Because there is more to anger or joy than the facial gestures, vocal cues, and the other factors we initially use to identify them in our pedestrian dealings, we are capable of misidentifying people's emotional states. The same is true of patterns of rational behavior. If Gabriel is a pathological individual who delights in deceiving people for its own sake, then his movements and utterances might lead me to mistake his will to deceive for a fear of being robbed (Section 8.4). In these cases, we mistake other people's thoughts, feelings, and perceptions because we fail to perceive the full range of factors that contribute to their behavior, just as we fail to see the mating combination in Figure 13.5.

Yet, according to hylomorphists, there are limits on how wrong we can be about other people's thoughts, feelings, and perceptions. We cannot be mistaken about them in the radical way that motivates the problem of other minds. The analogy with water is once again helpful for understanding this point. Misjudging that something is water depends on being able to identify the features typically associated with water.

Figure 13.5 Occluded pattern on a chessboard.



The blotch on the chessboard makes it impossible to discern whether there is a checkmate pattern.

When I judge that the liquid in this glass is water, I am identifying and responding to features of the liquid such as its look and taste. If a liquid looks like water, tastes like water, and has other features that I typically use to identify water, then I judge that the liquid is water. Importantly, I identify and respond to these features both when my judgment is accurate and when it is inaccurate—both when the liquid really is water and when it is something else. In both cases, my judgment is based on a prior ability to identify and respond to the features typically associated with water. Something analogous is true of thoughts, feelings, and perceptions on the hylomorphic view. Consider an example.

Paul Ekman and his colleagues have shown that various movements, facial expressions, and vocal cues manifest themselves involuntarily when we experience emotional states such as anger, joy, and fear. The *orbicularis oculi* muscle surrounding the eye, for instance, contracts involuntarily when people experience enjoyment. That same contraction is extremely difficult to produce voluntarily; in fact, Ekman (2007: 206) estimates that only a small percentage of people (about 10 per cent) can contract it voluntarily at all. (Actors who look convincingly as if they are enjoying themselves are most likely remembering something enjoyable, says Ekman, and the remembered emotion is responsible for producing the involuntary muscular contraction.) Something analogous is true of other psychological activities on the hylomorphic view. Gabriel's fear of being robbed by his neighbor might be expressed in a number of different ways—locking the door on short trips, setting up a security camera in his apartment, saying "I'm afraid my neighbor will rob me," or feeling uncomfortable at the prospect of leaving his apartment unlocked. But there are limits to the variability of fear behavior. If Gabriel never locked his door on short trips, never took security precautions of other sorts, denied that he was afraid of being robbed by his neighbor, or never felt uncomfortable with the idea of leaving his apartment unlocked, it would be doubtful that he really was afraid of being robbed by his neighbor. It is only because human psychological activities follow regular patterns like these that there is a possibility of misidentifying what thoughts, feelings, and perceptions people have.

When I judge that Eleanor is experiencing joy, I am identifying and responding to certain features of her behavior—the contraction of the *orbicularis oculi* muscle, for instance. Eleanor's face looks a certain way, and hence I judge that she is joyful. I identify and respond to that look both when my judgment is accurate and when it is inaccurate—both when Eleanor really is experiencing joy and when she is convincingly feigning it. In both cases, my judgment is based on a prior ability to identify and respond to the features associated with joy. Likewise, when I judge that Gabriel is afraid of being robbed, I am identifying and responding to features of his behavior such as his assertion that he is afraid of being robbed. His behavior looks and sounds a certain way, and on the basis of the way it looks and sounds, I judge that he is afraid of being robbed. I identify and respond to these features of Gabriel's behavior both when my judgment is accurate and when it is inaccurate. Making erroneous judgments about people's thoughts, feelings, and perceptions depends on the prior ability

to identify and respond to the features typically associated with those thoughts, feelings, and perceptions.

The upshot, say hylomorphists, is that we can be wrong about people's thoughts, feelings, and perceptions some of the time, but we cannot be wrong about them all of the time. If that were the case—if there were no social, environmental, or physiological features typically associated with this or that thought, feeling, or perception, or if we were not able to identify and respond to those features—then we would not be able to make judgments about other people's thoughts, feelings, and perceptions at all, rightly or wrongly. The possibility of us being wrong about other people's thoughts, feelings, and perceptions some of the time depends on us being right about other people's thoughts, feelings, and perceptions most of the time. According to hylomorphists, this is a further reason to think there can be no real problem of other minds. We must have a reliable source of information about other people's thoughts, feelings, and perceptions in order to be in error about what thoughts, feelings, and perceptions other people have. This implies two things. First, it implies that the second assumption motivating the problem of other minds is false. There is a limit on the range of gestures, movements, vocal cues, and other factors with which various thoughts, feelings, and perceptions can be correlated. I can use a grimace in an attempt to deceive you into believing that I am in pain only because grimaces are typically due to genuine feelings of pain. Second, it implies that even if my knowledge of other people's thoughts, feelings, and perceptions is fallible, it is not fallible in the radical way that motivates the problem of other minds. I cannot entertain serious doubts that other people think, feel, and perceive, for entertaining such doubts presupposes a reliable ability to identify thoughts, feelings, and perceptions.

13.7 Worry 8: Behaviorism, Dennett, and the Psychophysical Identity Theory

The hylomorphic approach to the problem of other minds is liable to remind some readers of logical behaviorism, and others of Dennett's (1991) real patterns. Hylomorphism and behaviorism, for instance, both reject the idea that thoughts, feelings, and perceptions are occurrences in an inner domain. Both deny that thoughts, feelings, and perceptions are hypothetical postulates of a prescientific theory, and both claim that thoughts, feelings, and perceptions are in some sense directly observable under pedestrian circumstances. Moreover, my earlier talk of patterns of social and environmental interaction might suggest to some readers that hylomorphism is like Dennett's view. Yet, despite these similarities, hylomorphism differs from logical behaviorism and Dennett's view in significant ways (Jaworski 2011: 338–9; 2012).

First, hylomorphism is incompatible with physicalism, behaviorism and Dennett's view are not. A commitment to physicalism is in fact part of what motivates behaviorists to analyze psychological expressions as abbreviated physical descriptions

of actual and potential behavior. Hylomorphists, by contrast, reject physicalism as part of their broader commitment to structures in nature (Section 12.1).

Second, hylomorphists do not conceive of behavior as narrowly as behaviorists do. Behaviorists tend to conceive of behavior in terms of bodily movements or utterances—occurrences that can in principle be exhaustively described by physics and that are observable under pedestrian circumstances (Jaworski 2011: 105–6). According to hylomorphists, however, behavior comprises more than this. Thoughts, feelings, perceptions, and actions all involve social and environmental factors in addition to physiological ones. Hylomorphists thus reject the behaviorist project of analyzing psychological predicates and terms into longer descriptions of actual and potential bodily movements and utterances. Psychological language instead describes distinctive patterns of social and environmental interaction—structured manifestations of powers that cannot be analyzed or reduced to unstructured bodily movements, physiological states, or dispositions.

Third, hylomorphists and behaviorists endorse different semantics for psychological expressions. According to behaviorists, psychological expressions operate like abbreviations for longer physical descriptions of bodily movements and utterances. According to hylomorphists, by contrast, psychological expressions operate like natural kind terms that refer to coordinated manifestations of the powers of our parts and surrounding materials.

In addition, hylomorphists reject Dennett's claim that psychological discourse is merely a framework for predicting and explaining physical processes in a way that is more efficient, if less accurate, than physics. For hylomorphists, psychological discourse and the special sciences in general are conceptual frameworks for describing and explaining behavioral structures at levels higher than those postulated by physics. We postulate patterns not simply for predictive and explanatory convenience, but because there really are higher-level patterns, and we are interested in describing and explaining what there really is.

Hylomorphists are not completely unsympathetic to some of the claims behaviorists and Dennett make. Behaviorists see correctly, they say, that psychological expressions imply bodily conditions, and they see correctly that human patterns of social and environmental interaction are complex enough that we can only describe some of them using counterfactual conditionals such as, "If the child were to ask the father for the candy, then the father would probably give the candy to the child." But, say hylomorphists, from these observations behaviorists draw the wrong conclusion. They conclude falsely that psychological predicates and terms are abbreviations for more complex physical descriptions of bodily movements and utterances instead of seeing that psychological predicates and terms express broad patterns of social and environmental interaction, and that bodily movements and utterances are merely contributing factors. The bodily movements and utterances that constitute what behaviorists think of as behavior represent only two of several types of conditions involved in patterns of behavior as hylomorphists conceive of them.

Likewise, Dennett sees rightly that beliefs and desires are not inner states of some sort, and he sees rightly that psychological description and explanation has an interpretive dimension of the sort that was evident in the example of Gabriel locking the door (Section 8.4), but Dennett wrongly supposes that psychological discourse is merely an instrument for making more efficient predictions about human behavior. According to hylomorphists, psychological discourse expresses structured manifestations of our powers and those of surrounding materials. To use a slogan: real patterns are more real for hylomorphists than they are for Dennett. Hylomorphists are committed to structure realism, whereas Dennett is committed to nonreductive structure antirealism (Section 1.2).

With the foregoing points in mind, we are in a position to address Williams' Worry 8. It claims that hylomorphism is committed to some type of reductive physicalism, either behaviorism or the psychophysical identity theory. The reason, it says, is that hylomorphists are committed to thoughts, feelings, and perceptions being coordinated manifestations of the powers of our parts and surrounding materials, and they are committed to at least some of these manifestations being directly perceivable in pedestrian circumstances. The idea that we can directly perceive other people's thoughts, feelings, and perceptions sounds like behaviorism, says the argument. Hylomorphists might reply that there is more to a thought, feeling, or perception than the coordinated manifestation of powers we can directly perceive, such as the manifestations of powers that are possessed by parts of our nervous systems. These are among the parts that manifest their powers when people remember, imagine, fantasize, or think "in their heads," without revealing their remembering, imagining, fantasizing, or thinking to others. Yet, the claim that our private remembrances, imaginings, fantasies, and thoughts are coordinated manifestations of the powers of parts of our nervous systems can sound like a kind of psychophysical identity theory. Hylomorphism thus appears to be committed to something like behaviorism when it comes to some mental states, and it appears to be committed to something like the psychophysical identity theory when it comes to others. Either way, says the argument, hylomorphism seems to be committed to something like reductive physicalism because both behaviorism and the psychophysical identity theory imply that physical discourse can in principle take over the descriptive and explanatory roles that psychological discourse plays.

Hylomorphists respond that Worry 8 is based on two false premises. The first claims that behaviorism is implied by a view committed to thoughts, feelings, and perceptions involving coordinated manifestations of powers that are directly perceivable in our pedestrian dealings. We can see that this premise is false in several ways. First, behaviorism claims that psychological descriptions are abbreviations for physical descriptions of actual and potential movements and utterances. The claim that thinking, feeling, and perceiving involve the coordinated manifestations of powers that are directly perceivable does not imply this semantic thesis. In fact, we have seen that hylomorphists endorse a completely different semantics for

psychological language (Section 13.6). Second, according to hylomorphists, there is more to thinking, feeling, and perceiving than the coordinated manifestation of the powers of things we can directly perceive. Thinking, feeling, and perceiving also involve the coordinated manifestations of the powers of things we cannot directly perceive, such as parts of the nervous system whose activities we can only discern using specialized equipment. Finally, there is the coordination of the manifestations—the activity-making structure that confers unity on them (Sections 8.1–8.2). This is not something included within the behaviorist framework. Among other things, the movements and utterances in terms of which behaviorists seek to analyze psychological descriptions are supposed to be describable using only the conceptual resources of physics, but structures, according to hylomorphists, including the kinds of structures that make thoughts, feelings, and perceptions, are not like this (Sections 6.3, 8.1). As a result, they reject physicalism (Section 12.1), unlike behaviorists, whose analyses of psychological expressions are motivated by a commitment to it. In addition, behaviorists have difficulty accommodating the distinction between movements and utterances, on the one hand, and intentional actions, on the other. This distinction fits comfortably within a hylomorphic framework since it claims that movements, utterances, and other physiological occurrences must be structured if they are to compose actions (Sections 13.3–13.4).

The second false premise on which Worry 8 depends is that the psychophysical identity theory is implied by a view that is committed to thoughts, feelings, and perceptions involving coordinated manifestations of powers that are *not* directly perceivable in pedestrian circumstances. We can see that this premise too is false. The psychophysical identity theory implies that thoughts, feelings, and perceptions are identical to internal states that can be exhaustively described in principle using only the conceptual resources of physics. But again, hylomorphists insist that there is more to thoughts, feelings, and perceptions than this: there is also the way various internal states are coordinated, the structures that make the activities of thinking, feeling, and perceiving what they are. For the reasons just mentioned, hylomorphists claim that these structures cannot be exhaustively described by physics.

The foregoing points are sufficient to dispose of Worry 8. A critic might nevertheless persist: admittedly, hylomorphism does not imply behaviorism or the psychophysical identity theory, yet it seems similar enough to one or the other that the objections levied against them can also be levied against it. To respond to this objection, it is sufficient to explain why hylomorphism is not vulnerable to the arguments typically advanced against behaviorism and the psychophysical identity theory.

There are at least three well-rehearsed objections to behaviorism:

(1) There are counterexamples to behavioral analyses of psychological expressions (Putnam 1963). It is possible that a super actor could replicate all the behavior associated with pain without actually being in pain. This shows that the movements and utterances typically correlated with pain are not sufficient

for being in pain. Nor are those movements and utterances necessary. It is possible that a super Spartan might experience pain without manifesting any of the behavior typically associated with pain—without even having dispositions to pain behavior. Likewise, people with locked-in syndrome can experience pain while being incapable of manifesting it.

(2) It is impossible even to formulate behavioral analyses of psychological expressions. Behaviorism implies that a statement about someone's psychological state can be translated into a statement about his or her behavior, and it implies that the behavioral conditions cited in that translation are sufficient for knowing that the original statement is true. But behavioral conditions are sufficient for knowing someone's psychological state only if we make countless further assumptions about that person's other psychological states (Chisholm 1957: chapter 11; Geach 1967: chapter 1).

(3) Behaviorism is committed to an implausible account of first-person authority (Ziff 1958). If psychological expressions are just abbreviations for descriptions of actual and potential movements and utterances which are directly observable under pedestrian circumstances, then it seems that I know about my thoughts, feelings, and perceptions by observing my own movements and utterances in more or less the way I observe the movements and utterances of other people. I know that I am in pain, for instance, by observing that my pupils are dilated, that I am sweating, wincing, and so on. But this seems implausible; surely I know that I am in pain by directly feeling it.

None of the foregoing objections applies to hylomorphism. Objections 1 and 2 target a theory that is committed to the possibility of analyzing psychological expressions into descriptions of actual and potential movements and utterances, but hylomorphism denies the possibility of any such analyses for reasons discussed earlier. Objection 3, for its part, applies to a theory that denies that we enjoy any special first-person knowledge of our own thoughts, feelings, and perceptions. But hylomorphism does not deny this. It claims that thoughts, feelings, and perceptions are coordinated manifestations of the powers of our parts and surrounding materials, and there is nothing about this claim that implies that I can know about those manifestations only in the way other people do. Nothing in the hylomorphic claim rules out the possibility that I could know about the manifestations of the powers of, say, the parts of my own nervous system in a way that other people cannot. Consequently, nothing prevents hylomorphists from adopting a higher-order perception theory (Armstrong 1968; Lycan 1996) or a higher-order thought theory (Carruthers 2000; Rosenthal 2005) to account for the knowledge we have of our own thoughts, feelings, and perceptions. A view of this sort would make two claims:

(1) We have the power to register directly the coordinated manifestations of the powers of some of our parts, such as those embodying our powers to think, feel, and perceive.

(2) Other people do not have the same power: they cannot register directly the coordinated manifestations of the powers of our own parts in pedestrian circumstances, and even if they are able to register the coordinated manifestations of the powers of those parts using specialized equipment, the way they register those manifestations is different from the way each of us does when exercising the power described by (1).

A view committed to (1) and (2) implies that each of us has a distinctive first-personal way of registering the activities of our nervous systems, which compose our own thoughts, feelings, and perceptions. Recall that on the hylomorphic view, we have the power to directly perceive at least some activity-making structures (Section 13.6). In line with that idea, a view committed to (1) claims that I have the power to register directly the coordinated manifestations of the powers of those parts of my nervous system which embody my capacities for thinking, feeling, and perceiving. According to (2), however, other people do not have the same power: they cannot directly register the manifestations of the powers of the parts of my nervous system. Even if someone could use specialized equipment, such as an fMRI, to know that, say, the cells in my limbic lobe were activated, this way of registering my limbic lobe activity would be different from the way I register that activity in my own case—a way that involves a power I am able to exercise in my pedestrian interactions without the need of specialized equipment.

It is worth mentioning that even though this view is compatible with us enjoying a distinctive type of first-person knowledge, it does not imply that this knowledge is infallible or incorrigible. We can err in our judgments of what we think, feel, and perceive in the first-person case, just as we can in the third-person case (Section 13.6). One reason for this is that our thoughts, feelings, and perceptions involve more than the physiological factors we can directly register using the power described in (1). They involve social and environmental factors that we might not be able fully to appreciate. I might not know what I think or how I feel about a new situation at work, about something my doctor has told me about my health, or about the current status of my marriage or family life. It is possible, moreover, that someone with more knowledge and experience could be better able to discern what I think or how I feel than I myself can, just as a chess master might be able to discern features of my own board position more accurately than I can.

The foregoing remarks should make it clear why Objection 3 does not apply to the hylomorphic view. Hylomorphism is compatible with each of us enjoying a distinctive type of first-personal access to the activities of our nervous systems—activities that in part compose our thoughts, feelings, and perceptions. Hylomorphism is not committed, therefore, to the implausible account of first-person authority that behaviorism is alleged to imply.

To complete the response to Worry 8, we still have to consider whether hylomorphism is liable to objections that have been levied against the psychophysical

identity theory. The best-known objection to the latter is the multiple realizability argument, the topic of Section 13.8.

13.8 Worry 9: Multiple Realizability

This section is concerned primarily with Worry 9, although it has implications for Worry 8 as well. Worry 9 claims that hylomorphism must imply some type of reductivism because it is incompatible with a central tenet of the antireductionist consensus, namely the multiple realizability thesis, the claim that a given type of psychological state, such as anger, fear, or desire, can be correlated with many different types of physical states—one type of physical state in humans, a different type in, say, extraterrestrials or robots. Why suppose hylomorphism is incompatible with multiple realizability? Hylomorphists claim that our thoughts, feelings, perceptions, and actions are essentially embodied in the kinds of parts we possess (Sections 8.2–8.3). The activities of those parts and surrounding materials constitute material conditions that, according to hylomorphists, are metaphysically necessary for us to engage in higher-level activities (Sections 8.2–8.3). Hylomorphists, moreover, appear to construe those material conditions so narrowly that they rule out the standard examples of multiple realizability. For example, hylomorphists would deny that a human could have all of its parts replaced by nonbiological prostheses (Section 6.4). Moreover, since the activities in which a human engages are composed partly by the activities of its parts (Chapter 8), a human who underwent the replacement of a substantial number of its biological parts with nonbiological prostheses would no longer be capable of engaging in a range of its typical activities—including perhaps thinking, feeling, perceiving, and acting. Given reasonable assumptions, this appears to imply that a robotic being could never think, feel, perceive, or act. Moreover, if the predicates and terms we use to refer to our thoughts, feelings, perceptions, and actions refer to activities that are essentially embodied in the kinds of parts we possess, then there is reason to think that according to hylomorphism, extraterrestrials that differ from us physiologically could never think, feel, perceive, or act. In fact, hylomorphism might even imply that familiar nonhuman animals such as dogs and cats cannot think, feel, perceive, or act. Some of these implications are implausible on their face, and they also contradict the multiple realizability thesis. Since that thesis has been the mainstay of antireductionist thinking since the late 1960s, it seems reasonable to conclude that hylomorphism must be committed to some type of reductivism.

Hylomorphists have a twofold response to Worry 9. First, they argue that antireductionism does not depend on the multiple realizability thesis. It can be based on something else, namely the hylomorphic notion of structure. Recall that according to hylomorphists, even though the activities of structured individuals are exhaustively decomposable into the manifestations of the powers of my parts and surrounding materials, there is still more to those activities than that: there is also the way those manifestations are structured or coordinated (Sections 8.1–8.2, 12.3). In order for me

to throw a baseball, the powers of my parts and surrounding materials cannot manifest themselves in just any old way: the right kind of coordination is essential to them composing the activity they do. That coordination is not a property of the parts and materials or their powers taken on their own; it is rather a property of the activity as a whole—a structure I impose on them. As a result, it is not possible to exhaustively describe and explain the activities of a structured individual simply by appeal to its parts and surrounding materials. Lower-level disciplines cannot take over the descriptive and explanatory jobs of higher-level ones, and so higher-level disciplines are irreducible to lower-level ones. William Bechtel makes a similar point:

> [T]he independence of levels does not require appeal to the multiple realizability of higher levels. There may only be one kind of mechanism that is fully appropriate for engagement in a particular causal process . . . Even so, different inquiries are needed to determine the variables that affect the behavior of the mechanism as a whole and to understand how the organized set of parts and operations enables the mechanism to behave that way. (2008: 155–6)

According to hylomorphists, antireductionism does not depend on there being one-many or many-many relations between higher-level and lower-level properties. Even if hylomorphism was not able to accommodate multiple realizability, it would still imply a robust antireductionism.

In fact, however, hylomorphism is able to accommodate multiple realizability, at least the basic idea behind it (Jaworski 2011: 341–2). This is the second aspect of the response to Worry 8. Hylomorphists agree that we can use psychological predicates and terms to describe and explain the behavior of familiar nonhuman animals such as dogs and cats. They can also admit that we can conceive of using those predicates and terms to describe and explain the behavior of alien species and perhaps even complex robotic systems. According to hylomorphists, however, our ability to apply psychological predicates and terms to these things does not depend on those predicates and terms having abstract definitions as functionalists claim. According to functionalists, mental properties are higher-order properties, ones whose definitions quantify over lower-order properties. The property of being in pain, for instance, is defined as the property of having some property with instances that are typically caused by certain environmental inputs and that typically cause certain behavioral outputs. If this is what pain is, then it is possible to apply the predicate 'is in pain' to things that differ from each other physiologically in a wide variety of ways. If dogs, cats, extraterrestrials, robots, and humans all have properties that correlate the right kinds of environmental causes and behavioral effects, then those things can be said to be in pain. According to functionalists, moreover, if that predicate expresses a genuine property—a higher-order property—then one and the same property can be instantiated by dogs, cats, extraterrestrials, robots, and humans.

We argued in Section 5.3, however, that there are no higher-order properties. If that is the case, then it rules out a functionalist analysis of mental properties. Does it

also rule out the possibility of applying the predicate 'is in pain' to dogs, cats, robots, and extraterrestrials? Hylomorphists insist not. The reason is that the application of psychological predicates to nonhumans can be based on something other than a functionalist analysis of psychological language. In particular, it can be based on analogy.

Hylomorphists claim that we can apply our human-defined psychological predicates and terms to nonhuman things by drawing analogies between their behavior and ours. Consider by analogy the way we can use the term 'punch' to describe the activity of someone with a prosthetic limb or a device that is not attached in any way to a human operator, such as a mechanized hole punch. Even if punching—strictly speaking—cannot be performed without a human hand oriented and moved in a specific kind of way, it might still be possible to apply our hand-defined term to something other than a hand by drawing an analogy between what it does and what a hand does. According to hylomorphists, something analogous is true of the way we apply psychological predicates and terms to nonhuman things. We can do so not because those predicates and terms are abstractly defined, as functionalists would have it, but because we have the ability to draw analogies.

Alva Noë describes a similar idea in connection with his enactive account of perception. He rejects the multiple realizability thesis as functionalists understand it, yet looks to accommodate something like it:

> To understand the nature of vision, as distinct from touch or hearing, you must focus on the differences in the respective patterns of sensorimotor dependence. Nevertheless . . . sensorimotor dependencies are themselves determined by the lower-level details of the physical systems on which our sensory systems depend. The eye and the visual parts of the brain form a most subtle instrument indeed, and thanks to this instrument, sensory stimulation varies in response to movement in precise ways. To see *as we do*, you must then have a sensory organ and a body like ours. This raises the worry about sensorimotor chauvinism . . . In virtue of *what* can the visual systems of the human, the crab, and the bumblebee all be deemed *visual* systems? . . . According to the enactive view, there is sufficient high-level, gross sensorimotor isomorphism between these different perceptual systems to count one and all as *visual* systems. (2004: 112–13)

According to Noë, various high-level, gross isomorphisms make it possible for us to catalog the subsystems of physiologically different organisms under a single heading. Likewise, on the hylomorphic view, there are similarities among the ways in which, say, reflected light factors into the behavior of physiologically diverse kinds of organisms. These similarities enable us to draw analogies between the behavior of those organisms and our behavior. Reflected light factors into the behavior of crabs and bumblebees in a way that is similar in some respects to the way light factors into our behavior. These similarities enable us to say that humans, crabs, and bumblebees can all see, and they enable us to identify which of their parts embody their respective capacities to see.

In the case of familiar nonhuman animals such as dogs and cats, drawing analogies between their behavior and ours is relatively easy. We have shared evolutionary histories that have given rise to similar body plans and biofunctional parts—heads, eyes, noses, mouths, and so on. These body plans make it easy for us to discern patterns in their behavior that are similar to the patterns in ours, and these similarities enable us to describe and explain their behavior using predicates and terms such as 'eats,' 'sleeps,' 'sees,' 'believes,' and 'wants.'[10] For similar reasons, it is conceivable that we might encounter alien species or even complex robotic systems whose behavioral patterns were similar enough to our own that we would be able to apply psychological predicates and terms to them as well. We could even conceive of applying those predicates and terms to alien species that differed from us radically in terms of their overall body plans and functional parts. Consider, for instance, a movie like *The Blob,* in which we are invited to describe and explain the behavior of the blob-like creature psychologically. Even though its body plan is very different from ours, the filmmakers include enough scenes displaying how it behaves to enable us to discern human-like patterns of desire or intention. If we were to encounter blob-like creatures in reality, and were able to observe enough of their behavior to discern behavioral patterns similar to those of humans, we might also be able to apply psychological predicates and terms to them as well.

The case of robotic systems is more problematic. The reason is that those systems are artifacts, and it is an implication of the hylomorphic view that artifacts do not exist: what we call artifacts are just physical materials spatially arranged artifact-wise (Section 6.4). Hylomorphists are nevertheless open to the empirical possibility that a robotic system could be a living thing, albeit an artificially created one. If that were the case—if a robotic system displayed distinctive activities which couldn't be accounted for by appeal to the properties of physical materials by themselves—then this would give us good reason to think that the system was a distinctive individual, like a living thing, only robotic (Section 7.4). Provided we could discern in the system's behavior ways of interacting with other things and the environment like ours, the system would be a candidate for psychological description. Hylomorphists are thus open to the empirical possibility that we might be able to describe and explain the behavior of extraterrestrials and complex robotic systems using a psychological vocabulary.

The foregoing points are sufficient to dispense with Worry 9. They are also sufficient to dispense with the remainder of Worry 8: they make it clear why the multiple realizability argument does not pose a threat to hylomorphism, as it does to the psychophysical identity theory.

[10] Wittgenstein seemed to have something like this in mind when he said "only of a living human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind; hears; is deaf, is conscious or unconscious" (1953, section 281).

13.9 Conclusion

This chapter described how hylomorphists deal with three mind-body problems: the problem of emergence, the problem of downward causation, and the problem of other minds. The problem of emergence is the problem of explaining how higher-level phenomena—including higher-level structural phenomena of the sort hylomorphists posit—are produced or generated by lower-level physical processes. Hylomorphists argue that requesting such an explanation is misguided. It assumes that higher-level structural phenomena are produced or generated by lower-level phenomena. But this is an assumption hylomorphists reject. Structures are not in general produced by the lower-level things they structure. The arrangement of wood and metal that makes a piano is not produced by the wood and metal; it is instead an order imposed on them. It is thus illegitimate to request an explanation of how, say, brains produce consciousness, for such a request assumes there is a kind of occurrence that hylomorphists deny exists, namely the generation of higher-level phenomena by lower-level occurrences. Requesting an explanation of such a thing would be as illegitimate from a hylomorphic perspective as requesting that a meteorologist explain how the will of Zeus produces rain.

The problem of downward causation is the problem of explaining how higher-level phenomena are able to make a causal difference in the world. The problem of mental causation is one species of the problem of downward causation. It can be formulated in terms of five jointly inconsistent claims: (1) Actions have mental causes; (2) actions have physical causes; (3) the mental and physical causes of actions are distinct; (4) if actions have multiple causes, then they are overdetermined; and (5) actions are not overdetermined. The hylomorphic solution to the problem is based on the causal pluralism described in Chapter 10: there are many different kinds of causes and many different kinds of causal relations. There is, in particular, a difference between the way thoughts, feelings, and perceptions cause or explain actions, and the way physiological mechanisms do. The former rationalize actions; the latter contribute to actions as subsystems. Because of this distinction among causes, claims (1)–(3) must be rewritten: (1′) actions are rationalized by thoughts, feelings, and/or perceptions; (2′) muscular contractions are triggered by events in the nervous system; and (3′) rationalizing causes and physiological triggers are distinct. In addition, claim (4) is false: explanations that appeal to reasons and explanations that appeal to physiological mechanisms pick out causal factors of different sorts; they answer requests for different kinds of information, just as explanations that appeal to roadway grading and explanations that appeal to blood-alcohol levels provide different kinds of information about car crashes. There is thus no threat of actions being overdetermined by reasons and physiological triggers. Finally, hylomorphists say that claim (5) must be rewritten in light of the triggering/rationalizing distinction. Here are two possible ways of doing so: (5′) An action does not have more than one rationalizing cause; (5″) an action does not have more than one

physiological trigger. Both of these claims and also their denials are compatible with (1′)–(3′), so whether hylomorphists accept both claims or deny either, they succeed in solving the problem.

Hylomorphism also has implications for the problem of other minds. That problem arises on account of two crucial claims. First, we can know about other people's thoughts, feelings, and perceptions only by making inferences from bodily behavior. Second, the same bodily behavior can be correlated with many different kinds of thoughts, feelings, and perceptions. If these claims are true, then someone's bodily behavior does not provide decisive evidence that he or she has this or that thought, feeling, or perception; we thus have difficulty knowing what other people's thoughts, feelings, and perceptions are, or perhaps whether they have any thoughts, feelings, or perceptions at all. According to hylomorphists, however, both assumptions are false. They claim that thoughts, feelings, and perceptions are coordinated manifestations of the powers of our parts and surrounding materials, that we can perceive some of those coordinated manifestations, and that some of the coordinated manifestations we can perceive fix the referents of the predicates and terms we use to describe thoughts, feelings, and perceptions. What terms like 'anger,' 'joy,' and 'fear' refer to on this account are structured activities, coordinated manifestations of the powers that people and various other things have. According to hylomorphists, this is what thoughts, feelings, and perceptions are: they are phenomena of the same sort as throwing a baseball, coordinated manifestations of the powers of our parts and surrounding materials.

This understanding of psychological language denies that psychological explanation is a species of theoretical explanation. Psychological predicates and terms do not refer to or express unobserved hypothetical entities that are posited to explain directly observable human behavior; they instead refer to or express directly observable manifestations of the powers of individuals and surrounding materials—what I have sometimes referred to as 'patterns of social and environmental interaction.' The way we perceive these patterns is analogous, I would suggest, to the way we perceive patterns in chess or music. Empirical work like psychologist Paul Ekman's supports the idea that normal humans are equipped to discern emotions, intentions, and other psychological patterns in each other's behavior, and if that is true, then the first assumption motivating the problem of other minds is false. We directly perceive that other people are amused, joyful, or afraid. Importantly, the hylomorphic view does not imply that our knowledge of other people's thoughts, feelings, and perceptions is infallible. It looks to accommodate the kind of fallibility we experience in our ordinary dealings with one another while at the same time rejecting the more radical fallibility that motivates the problem of other minds. The hylomorphic approach to the problem of other minds is liable to remind some readers of logical behaviorism, and others of Dennett's (1991) real patterns. But hylomorphism differs from these other views in important ways. These differences, moreover, provide a basis for addressing Williams' Worry 8, which is based on the premise that hylomorphism

is committed either to some type of behaviorism or to some type of psychophysical identity theory.

Finally, hylomorphists are able to deal with Williams' Worry 9, which claims that hylomorphism must imply some type of reductivism because it is incompatible with the multiple realizability thesis. Hylomorphists' response to the worry is twofold. First, they deny that antireductionism depends on the multiple realizability thesis; antireductionism can instead be based on a commitment to hylomorphic structure. Second, they argue that they can accommodate multiple realizability intuitions, albeit in a way different from the way functionalists do. Even though our psychological predicates and terms are defined in a human-centered way, it is nevertheless possible to extend their use to nonhumans by drawing analogies between the behavior of those nonhumans and our behavior.

14

Why Hylomorphism?

14.1 The Intuitive Appeal of Hylomorphism

We've considered hylomorphism and its implications for mind-body problems. The goal of this chapter is to present some reasons for endorsing it. My goal is modest: to show that hylomorphism deserves a place at the table alongside more familiar theories such as nonreductive physicalism, emergentism, and Russellian monism.

We can start by recognizing that hylomorphism is able to solve a variety of philosophical problems in ways that do not violate our commonsense intuitions, or that do not violate them to any greater degree than its rivals. We've seen, for instance, that it is able to solve problems with composition such as the body-minus problem without claiming that people don't exist, or that brains or hearts don't exist, or that there are objects like the object composed of my right hand and the Eiffel tower, or that there is an object that shares all its parts with me but that isn't me (Section 7.1). On these points, I think the implications of hylomorphism correspond more closely to the deliverances of commonsense than the implications of rival views. The same is true, I think, of its solutions to mind-body problems. It solves the problem of downward causation without making thoughts and feelings epiphenomenal, without claiming that physical laws are violated every time we act, without claiming that actions are overdetermined, and without endorsing reductivism (Sections 13.3–13.4). It solves the problem of emergence, moreover, without taking psychophysical emergence to be a brute matter of fact, without attributing mental states to microphysical entities, and without having to explain how mental states emerge out of combinations of protomental states (Section 11.2). Finally, it solves the problem of other minds by making the possibility of doubting other people's mental states depend on knowing that they have mental states (Section 13.6). I think the elegant way hylomorphism dispatches these problems speaks in its favor as a view that deserves serious consideration.

Hylomorphism also jibes with commonsense intuitions about the existence of structure. These include the intuitions brought out by the squashing example (Section 1.1). When we place Godehard in a leak-proof bag and squash it with several tons of force, the contents of the bag undergo a change. The physical materials that were in the bag before squashing remain the same, but before squashing the contents of the bag included one human being, whereas after they include none.

What explains the difference? The answer seems to be that squashing destroyed the way the physical materials in the bag were structured, and that Godehard's existence and the existence of his capacities to think, feel, perceive, and act depended on that structure. Commonsense thus seems to accord to structure the status of an ontological and explanatory principle, one that concerns what things are, and what they can do.

Johnston (2006) notes a body of empirical work that supports the idea that a hylomorphic conception of things is highly intuitive. Prasada et al. have achieved experimental results that indicate that our intuitive notion of individual objects, in contrast to continuous stuffs, is based on a notion of structure:

> A simple, and we believe correct, way of describing what distinguishes our conception of an entity as an object of some kind rather than an amount of material is the manner in which we think about the entity's structure . . . the manner in which the constituents of an entity are organized . . . This difference can be characterized as follows:
>
> *Object construal*: To conceive of an entity as an object of some kind is to understand it to have the structure it does, and not some other structure, because it is the kind of thing it is. That is, its structure is nonarbitrary and could not be just any other way. Having that structure is partially constitutive of being that kind of thing.
>
> *Substance construal*: To conceive of an entity as an amount of stuff is to think that the entity does not have the structure it does because it is the kind of thing it is. One must look to something other than what it is to understand why it has the structure it does. Thought of as stuff, its structure is arbitrary. Having that structure is not even partially constitutive of being that kind of stuff. (2002: 143–4)

Hylomorphism also accommodates commonsense intuitions about the super physicist (Section 1.1). The super physicist, recall, has complete knowledge of all the fundamental physical entities in the universe, but lacks the perceptual and conceptual resources to distinguish living things from nonliving ones, or mental beings from nonmental ones. The concepts of life, perception, desire, belief, money, sex, and so on are completely beyond its ken. As a result, its descriptions of the universe are framed solely in terms of the characteristics of fundamental physical materials. Many people have the intuition that the super physicist's descriptions of the world are missing some very important things. Hylomorphism accommodates this intuition. It says that the super physicist's descriptions of the world are missing the various ways that physical materials are structured or organized in the natural world—ways that mark the difference between living things and nonliving ones, or mental beings and nonmental ones.

Hylomorphism also meshes with the intuitions many philosophers have about organization. John Heil, for instance, says that "the world presents us with endless levels of complexity and organization" (2003: 245). Since Heil explicitly rejects higher-order properties, it seems reasonable to conclude that the levels of organization he has in mind are not mere logical constructs. But if that is the case, it becomes

difficult to avoid the conclusion that when he talks about organization, he has in mind a real ontological principle—something like hylomorphic structure.[1]

David Armstrong employs a notion of organization similar to the hylomorphic one as well. We saw an example of this in Chapter 1: "a man is a physical object," he says, "distinguished from other physical objects only by the special complexity of his physical organization" (1968: 11). If there is a difference between Armstrong's notion of organization and the hylomorphic one, it is that Armstrong does not take the organization that characterizes living things to pose a challenge to physicalism. Hylomorphists, as we have seen, disagree (Section 12.1).

Philip Kitcher has also employed a notion of organization like the hylomorphic one:

> [T]o the extent that we can make sense of the present explanatory structure within biology—that division of the field into subfields corresponding to levels of organization in nature—we can also understand the antireductionist . . . claim that . . . the current division of biology [is] not simply . . . a temporary feature of our science stemming from our cognitive imperfections but [is] the reflection of levels of organization in nature. (1984: 369, 373)

The same is true of Nancy Cartwright, although her preferred terms are 'arrangement' and 'configuration.' Importantly, she situates this notion amidst several others which are familiar features of the hylomorphic view. One is causal pluralism:

> [T]here is a great variety of different kinds of causes and . . . even causes of the same kind can operate in different ways . . . One factor can contribute to the production or prevention of another in a great variety of ways. There are standing conditions, auxiliary conditions, precipitating conditions, agents, interventions, contraventions, modifications, contributory factors, enhancements, inhibitions, factors that raise the number of effects, factors that only raise the level, *etc* . . . . [T]he term 'cause' . . . is *unspecific*: being abstract, the term does not specify the form that the action in question takes. (Cartwright 1999: 104, 119)

Because there are so many different ways that causes can operate, she says, "[r]eliable tests for whether one factor causes another must be finely tuned to how the cause is supposed to function" (1999: 120). Scientists secure the finely-tuned test conditions they need by constructing what she calls 'nomological machines':

> [A nomological machine] has fixed components with known capacities arranged in a stable configuration. It is the fixed arrangement among the parts that guarantees the causal relations depicted in [graphs of causal structures] . . . [W]e always need a machine like this to get laws—any laws, causal or otherwise. Sometimes God supplies the arrangements—as in the planetary systems—but very often we must supply them ourselves. (1999: 122)

[1] The same is true of the way Chalmers sometimes speaks of organization (which can be contrasted with the way he often speaks of structure (Section 1.1)). "My functional organization," he says, is "the pattern of causal organization embodied in the mechanisms responsible for the production of my behavior" (1996: 97). Since Chalmers does not reject higher-order properties, however, it is less clear what the ontological status of patterns and of organization is on his view than it is on Heil's.

Two points are noteworthy here. First, Cartwright's use of the term 'structure' is different from the hylomorphic use. She follows Spirtes et al. (1993) in using the term to refer to abstract entities. Causal structures in their sense are ordered pairs <*V*,*E*>, where *V* is a set of variables, and *E* is a set of ordered pairs <*X*,*Y*> of *V* which belong to *E* if and only if *X* causes *Y*. Causal structures represent relatively stable relations among a variety of causal factors. Cartwright argues that the stability of those relations depends on the reliable operation of a nomological machine. What enables the machine to operate reliably, moreover, is the arrangement or configuration of its components:

> [C]ausal structure arises from a nomological machine and holds only conditional on the proper running of the machine . . . Fixed patterns of association among measurable quantities are a consequence of the repeated operation of factors that have stable capacities arranged in the 'right' way. (1999: 134, 138)

This brings us to the second point: Cartwright's notion of configuration or arrangement is like the hylomorphic notion of structure, and her nomological machines are like structured individuals in the hylomorphic sense—especially the individuals that, as she puts it, "God supplies" (1999: 122). The notion of a nomological machine does not correspond exactly to the hylomorphic notion of a structured individual. For one thing, the notion of a nomological machine is broader. The solar system is a nomological machine according to Cartwright, but we have seen that it is most likely not a structured individual according to hylomorphists (Section 6.3). Despite this difference, the notions of a structured individual and a nomological machine are similar enough that it is not implausible to suggest that hylomorphism provides a way of understanding the metaphysics implicit in Cartwright's philosophy of science.

The foregoing are just some examples that illustrate the intuitive appeal that hylomorphism ought to have for some philosophers and laypeople.

14.2 Structure Realism versus the Alternatives

One general difference between hylomorphists and many of the authors who appeal to notions of organization similar to the hylomorphic one is that hylomorphists take seriously the ontological demands that talk of organization seems to imply. Recall that the kind of hylomorphism I've defended is committed to ontological naturalism, the claim that when it comes to determining what exists, empirical investigation is our best guide (Section 1.2). Ontological naturalism implies that if our best empirical descriptions, explanations, and methods posit entities of kind *K*, then we have good prima facie reason to think that *K*s exist. Many of our best empirical descriptions, explanations, and methods appear to posit various kinds of structure, order, arrangement, organization, or configuration. Here are two examples:

> Neurons exchange information at specialized structures called synapses. This transfer . . . requires special subcellular structures and highly refined physiological mechanisms . . . The

cell membrane's lipid bilayer is a matrix within which a number of other structures, mainly glycoproteins . . . reside . . . Ion channels . . . are membrane-spanning proteins . . . [and] are particularly important in neurons. They are composed of long sequences of amino acids arranged into . . . helical domains . . . [T]hese domains aggregate to form a wedge-shaped structure . . . The cerebral cortex is composed of a 2- to 4-mm-thick superficial layer of gray matter with an underlying mass of white matter. The gray matter is divided into six horizontal layers . . . [A]ll neocortex exhibits this six-layer organization . . . Approximately 300 to 600 cells of the neocortex are interconnected vertically to form a local functional unit, a cortical column . . . The vertical organization of cells within a column is superimposed on the horizontal organization of the columns across the cerebral cortex . . . [W]ithin a single sensory column in area 3a, all of the cells respond only to a stimulus of a single modality from a specific location . . . An adjacent column responds to a different modality . . . from the same location . . . This pattern repeats in an organized manner across the cortex. (Kingsley 2000: 92, 108, 174, 177)

[B]ehavioral studies of many groups of living vertebrates have advanced to the point where the detailed morphology of unusual structures can be closely correlated with the utilization of these structures as intraspecific combat and display organs . . . As a result, it is now possible to apply general principles developed from behavioral studies of living species to the interpretation of the morphology of extinct species. (Hopson 1975: 22)

Ontological naturalism implies that descriptions like these give us good prima facie reason to think that structure, order, arrangement, organization, or configuration exists; they thus make a serious ontological demand. Structure realism is the most straightforward way of meeting this demand. It takes empirical appeals to structure at face value. Structure, it says, is a basic ontological and explanatory principle: descriptions, explanations, and methods that posit structure cannot in general be reduced to, or paraphrased, or eliminated in favor of nonstructural descriptions and explanations. This straightforward approach to structure is the one favored by hylomorphists, and the one that recommends itself to philosophers who are sympathetic to principles like ontological naturalism and who appeal to notions of organization or structure. For this reason alone, those philosophers ought to take hylomorphism seriously.

Structure realism, however, is not the only way of accommodating empirical appeals to structure like the foregoing. We have seen that there are at least three alternatives, each of which claims that it is possible in principle to describe and explain everything without appeal to structure (Section 1.2). Structure eliminativism claims that statements about structure are literally all false, that talk of structure is the byproduct of a defective way of trying to describe and explain the behavior of things. Structure reductivism claims that some statements about structure are true, but that what makes them true is that the structures they posit can be identified with things that can be exhaustively described and explained without appeal to structure; as a result, nonstructural discourse can take over all the descriptive and explanatory jobs that appeals to structure perform. Finally, nonreductive structure antirealists look to retain the sparse nonstructural ontology that motivates eliminativism and reductivism, but without denying the legitimacy of structural discourse and without carrying the empirical burden of identifying structures with nonstructural things. Whatever

descriptive and explanatory legitimacy structural discourse has is not grounded in the identification of structures with nonstructural things; it is instead grounded in our descriptive and explanatory interests—interests that nonstructural discourse cannot satisfy. It is nevertheless possible to paraphrase statements about structure in ways that minimize their ontological implications—ways that enable us to say what we want to say without positing structure. Structure eliminativism, structure reductivism, and nonreductive structure antirealism are alternative ways that philosophers sympathetic to ontological naturalism might look to accommodate the ontological implications of empirical appeals to structure. There are nevertheless reasons to prefer structure realism.

First, structure realism is the simplest, most direct explanation for why appeals to structure in biology, neuroscience, and other scientific disciplines are successful. Appeals to structure are successful, it says, because there really are structures. If there is a better explanation, structure realists say, the burden of proof is on structure's opponents to provide it. In the absence of compelling reasons to do otherwise, however, we should take talk of structure in the sciences at face value. Structure realism should thus operate as our default position. It is up to its opponents to advance compelling reasons to reject it.

Second, structure realism enjoys certain advantages over the alternatives. The reason for preferring structure realism to structure eliminativism is a familiar reason for preferring realism to antirealism in general. Putnam states it as follows: "The positive argument for realism is that it is the only philosophy that doesn't make the success of science a miracle" (1975d: 73). We appeal to structure in our scientific dealings very effectively, and this effectiveness requires some type of explanation. The most obvious explanation is that there really is such a thing as structure. If there is not, then it is unclear how the effectiveness of appeals to structure in the sciences can be explained at all. If there really is no such thing as, say, the structure of the brain, then how do descriptions and explanations that appeal to brain structure manage to enjoy any success at all? Eliminativists do not appear to have a good answer, but realists do: descriptions and explanations that appeal to brain structure manage to be successful because there really is such a thing as brain structure. Structure realism's ability to explain the success of empirical appeals to structure is a reason to prefer it to structure eliminativism.

There are also reasons to prefer structure realism to structure reductivism and nonreductive structure antirealism. For one thing, the latter alternatives have a much heavier evidential burden than structure realism. By analogy, consider one of the reasons Putnam gave for preferring functionalism to the psychophysical identity theory:

> [W]e identify organisms as in pain . . . on the basis of their *behavior*. But it is a truism that similarities in the behavior of two systems are at least a reason to suspect similarities in the functional organization of the two systems, and a much *weaker* reason to suspect similarities in the actual physical details. (1967: 436)

According to Putnam, any behavioral evidence that supports the identity theory also supports functionalism, and it supports functionalism more strongly since identity theorists have to supply additional evidence to support the psychophysical identities they postulate. Functionalists thus have a lighter burden of proof than identity theorists: it is evidentially easier to be a functionalist than an identity theorist.

For analogous reasons, it is evidentially easier to be a structure realist than a structure reductivist. Structure reductivists have to show that the structures posited in empirical descriptions and explanations can be identified with things that can be exhaustively described and explained without appeal to structure. Likewise, nonreductive structure antirealists have to show that all references to structure can be paraphrased in ways that have no implicit or explicit commitment to it. Both of these are very ambitious projects. Hylomorphists do not have a priori arguments that these projects cannot be carried out. From their perspective, the feasibility of these projects needs to be evaluated empirically. The important point, they insist, is that they needn't carry out any analogous project. Biologists, neuroscientists, and others successfully describe and explain things by appeal to structure, and the success of their descriptions and explanations does not depend on a further assumption that what they say can be reduced to or paraphrased in favor of nonstructural descriptions and explanations. If it is possible exhaustively to describe and explain things without appeal to structure, as structure realism's opponents claim, then the burden is on them to establish this. Reductivists must establish that nonstructural discourse can take over the descriptive and explanatory roles of structural discourse, and nonreductivists must establish that appeals to structure can always be paraphrased in nonstructural ways. To date, however, neither of these points has been established. Reductivists and nonreductivists must instead issue promissory notes. The advantage for structure realists is that they needn't issue any such notes. Because structure realists take empirical claims about structure at face value, they do not have the additional burden of showing that we can in principle do without descriptions and explanations that posit structure. It is thus easier evidentially to be a structure realist than either a structure reductivist or a nonreductive structure antirealist.

In addition, nonreductive structure antirealism appears to face a problem analogous to the problems facing structure eliminativism. According to nonreductive structure antirealists, we have many different descriptive and explanatory interests, and because of that, we tend to describe and explain things using many different vocabularies, including a structural vocabulary. By using structural predicates and terms, we satisfy descriptive and explanatory interests that we could not satisfy if we limited ourselves to nonstructural ones. Predicates and terms such as 'structure' and 'organized,' however, do not refer to or express any features of the world that cannot in principle be described and explained nonstructurally. When biologists speak of the structures of cells or the organization of the nervous system, they are not describing anything over and above unstructured processes, say nonreductive structure antirealists, they are simply describing unstructured processes in a different way.

Structure is not something written in the book of nature as it comes off the press; it represents instead notes we jot in the margins—our commentary on a text written entirely in nonstructural terms. If this is the case, however, then how do appeals to structure manage to enjoy any descriptive, predictive, or explanatory success if fundamentally everything is unstructured? Simply because appeals to structure satisfy some of our interests does not guarantee that those appeals will manage to be successful at describing, predicting, or explaining anything. Accounting for the descriptive, predictive, and explanatory success of structural discourse requires something in addition to its ability to satisfy our interests.

In addition, if nonreductive structure antirealists are right, and we use structural predicates and terms simply because they enable us to satisfy peculiar descriptive and explanatory interests, it seems reasonable to ask why we have the interests we do, and why structural discourse is able to satisfy them. Surely, say hylomorphists, the best explanation is that we are interested in describing and explaining the real behavior of things, and that behavior involves various kinds of structure. Consequently, the best explanation for the descriptive and explanatory interests we have—the very interests nonreductive structure antirealists appeal to—is that structure really does correspond to something deep in reality. Until these challenges are answered satisfactorily, say hylomorphists, there are good reasons to reject nonreductive structure antirealism in favor of the more direct realist alternative.

Consider now a general argument that opponents of structure realism might advance (Jaworski 2011: 301–2). It appeals to Ockham's razor. In general, says the argument, we should not multiply entities beyond necessity. Other things being equal, we should choose the most ontologically parsimonious theory we can. Consequently, if the behavior of bodies can be exhaustively described and explained in nonstructural terms, we should not seek to explain the behavior of bodies by appeal to structure as well. Hence, a view that rejects structure should be our default position.

Hylomorphists can respond to this argument in at least two ways. First, while it might be true that we should prefer a more parsimonious theory, all things being equal, all things might not be equal. Ontological parsimony only becomes a decisive factor in theory choice when competing theories are all coherent, all consistent with the empirical data, and equal in explanatory power. If theories that reject structure are not coherent, if they have insoluble philosophical problems, if they are not consistent with the scientific data, or lack the explanatory power of a theory that endorses structure, then it does not matter how parsimonious their ontologies are, these other factors trump their parsimony. We saw a moment ago, moreover, that according to hylomorphists, opponents of structure have yet to establish that their theories are consistent with the empirical data, and have explanatory power to rival a theory that endorses structure: they have yet to establish that we can describe and explain living behavior without appeal to structure, or that we can reduce or paraphrase appeals to structure. It remains unclear, therefore, whether an appeal to parsimony gains any traction here.

Second, hylomorphists can argue that a theory can have complexities that offset its greater ontological parsimony. There are different ways of reckoning a theory's simplicity. Counting the kinds of entities in a theory's ontology is one. Another concerns the system of predicates, terms, laws, and other bits of machinery that enable the theory to perform its descriptive and explanatory jobs. The complexity of these other things can offset the simplicity of its ontology. Consider the debate between realists and nominalists about universals. Nominalism has a simpler ontology since it comprises only particulars instead of particulars plus universals. Realists argue, however, that nominalists purchase this simplicity at a price: they have a complex and often unsystematic semantics for predicates and abstract nouns, and this semantic complexity is a theoretical liability that offsets nominalism's ontological simplicity. Structure realists can argue against their opponents in a similar way. Consider, for instance, how this kind of argument could be advanced against nonreductive structure antirealists.

First, there is a sense in which nonreductive structure antirealism does not have a simpler ontology than structure realism. It claims that talk of structure is successful because it satisfies our descriptive and explanatory interests. Interests thus play a role analogous to the role that structure itself plays in structure realism: interests and structures both explain the success of empirical appeals to structure. Second, it seems clear that nonreductive structure antirealism will be more complex than structure realism since any story that nonreductive structure antirealists tell about the success of structure talk will require a semantics for structure talk that is more complicated than the semantics of structure realism. As a result, hylomorphists can say, even if parsimony ends up cutting against structure in the end, there are other considerations that cut equally against nonreductive structure antirealism.

The foregoing considerations show that hylomorphism's commitment to structure realism is preferable to the alternatives. This provides a further reason to take hylomorphism seriously.

14.3 Hylomorphism versus Other Mind-Body Theories

We have seen that hylomorphism provides a distinctive way of understanding the place of thought, feeling, and perception in the natural world (Chapters 8–13). There are reasons to prefer this approach to the alternatives, including emergentism, nonreductive physicalism, and Russellian monism. Like hylomorphism, these theories are naturalistic: they claim that we are physical beings with physical components, and that our distinctive powers to think, feel, perceive, and act are essentially embodied in the physical materials that compose us. They are also all antireductive: they deny that the descriptions and explanations of our powers are reducible to the descriptions and explanations provided by physics, chemistry, and neuroscience. Among theories with this nonreductive naturalistic profile, there are reasons to prefer hylomorphism.

First, we've seen that hylomorphism enjoys advantages over emergentism because it solves the major problems emergentists face: the problem of emergence and the problem of downward causation (Sections 13.1–13.4). It solves the first problem by denying that higher-level phenomena are produced or generated by lower-level phenomena. It thereby undermines the legitimacy of asking for an explanation of how such generation is possible (Section 13.2). Hylomorphism solves the second problem by appeal to causal pluralism (Section 13.3). Actions are rationalized by thoughts, feelings, and/or perceptions. By contrast, the muscular contractions and other physiological occurrences that compose actions are triggered by events in the nervous system or the surrounding environment. Rationalizing causes and triggering causes are distinct, and they contribute to explanations of their effects in different ways: they play different, noncompeting causal roles. As a result, the occurrence of one kind of cause does not trump the contribution of the other, and their co-occurrence does not imply that human behavior is overdetermined.

Hylomorphism has an edge over nonreductive physicalism for similar reasons. Kim (1993b) has argued that nonreductive physicalists face a problem with downward causation analogous to the problem facing emergentists. If physicalism is true, then everything is physical, including any properties that could make a causal difference to things. Nonreductive physicalists look to deny that mental properties are physical properties; they typically do this by claiming that mental properties are (so-called) higher-order properties. We have already argued that there are no higher-order properties (Section 5.3), and if that is the case, nonreductive physicalism is a nonstarter. Even if there are higher-order properties, however, Kim argues that nonreductive physicalists face a problem. If mental properties are not identical to physical properties, then it looks like nonreductive physicalists must deny that mental properties play any role in the production or explanation of actions, for if physicalism is true, everything can be described and explained exhaustively by physics. If, on the other hand, nonreductive physicalists insist that mental properties play a role in the production or explanation of actions, then it looks like they are committed to claiming that mental and physical causes overdetermine their effects. They are, moreover, in danger of abandoning physicalism in favor of a property dualism that posits nonphysical mental properties. If they insist that mental properties are not in fact distinct from physical properties, then they are in danger of abandoning their antireductionism. Nonreductive physicalists are thus faced with several uncomfortable options. In the first case, their view ends up collapsing into a kind of eliminativism or epiphenomenalism; in the second, it ends up collapsing into a kind of property dualism; and in the third, it ends up collapsing into a kind of reductivism. There are ways nonreductive physicalists can respond to Kim's argument, but those ways of responding depend on rejecting metaphysical assumptions that there are good reasons to endorse (Jaworski 2011: 169–76). Insofar as a hylomorphic approach to downward causation avoids this tangle, it has an edge over nonreductive physicalism.

In fact, hylomorphism arguably has an edge over physicalist theories in general for at least two reasons. The first is analogous to the one that gives structure realism an edge over its reductive and antirealist rivals: it is evidentially easier to be a hylomorphist than it is to be a physicalist. Unlike physicalists, hylomorphists do not make a totalizing claim about the descriptive and explanatory power of physics, and they make no analogous claim about any other branch of science. Physicalists, by contrast, must look to establish that it is possible in principle for physics to describe the whole of reality. This is an astoundingly ambitious claim. Hylomorphism is much more moderate. It recognizes a plurality of structural levels and a corresponding plurality of explanatory and scientific kinds. It does not shoulder the burden of showing that the descriptions and explanations produced by the special sciences are all false, or that they are all reducible to or paraphrasable in terms of descriptions and explanations produced by physics. Physicalists thus have a much heavier burden of proof than hylomorphists, and this gives the latter an edge.

Hylomorphism enjoys a second advantage over physicalism: it avoids the problem of having to provide a clear definition of the term 'physical.' Physicalism claims that everything is physical. Recall, however, that Hempel's dilemma poses a problem for any attempt to define what 'physical' means (Section 11.1). For a physicalist, a workable definition of 'physical' must be continuous with the descriptions and explanations provided by empirical disciplines like physics. But those disciplines are progressive. The physics of the past is not the same as current physics, which will not be the same as the ideal physics of the future. If physics is defined relative to a preliminary stage in its development at which its theories are still subject to falsification and revision, then physicalism ends up being false since the physical theories advanced at that stage are false. If, on the other hand, physics is defined relative to a future, ideal stage of its development, then physicalism ends up lacking content. Since we do not yet know what the future, ideal physical theory says, we do not know what it means to say that everything is the way the final, ideal physical theory says it is. As a result, Hempel's dilemma concludes that physicalism is either false or else lacking in any content that we can evaluate. There are other ways of defining 'physical,' of course, but Montero (1999) has argued that most of these are no help to the physicalist. The problem generalizes, moreover, to any theory that trades on a clear definition of 'physical.' A DAT that claims that mental and physical properties are distinct is no better off than a physicalist one in this regard since the property dualism at its core presupposes that we know enough about what the physical is to make the distinctness of mental and physical properties interesting.

By contrast with theories of the foregoing sorts, hylomorphism does not depend on having a clear definition of 'physical.' We saw in Section 8.5 that nothing about the hylomorphic framework of powers and manifestations forces hylomorphists to adopt one definition of 'physical' as opposed to another, or to draw any distinction between the physical and nonphysical or the mental and nonmental. Hylomorphists are free to adopt definitions of 'mental' and 'physical' or to draw distinctions between

what is mental and what is physical as suits their descriptive and explanatory purposes. According to them, we can get on with our empirical investigation of powers, parts, and manifestations without ever employing those categories. As a result, hylomorphists do not face Hempel's dilemma or what Montero calls 'the body problem,' the more basic problem of providing a workable definition of 'physical' that renders mind-body debates intelligible.

Hylomorphism also has advantages over what Chalmers (1996; 2002a) calls 'Russellian monism.' Russellian monism claims that fundamental physical entities have categorical mental properties—either full-blown mental properties like the ones you and I have, or else protomental properties that combine in increasingly complex ways to produce full-blown mental properties. Mentality is thus present throughout the natural world, even at the fundamental level of reality (hence the labels 'panpsychism' and 'panprotopsychism'). The same categorical mental or protomental properties are nevertheless responsible for generating the dispositions that are the bases of our physical descriptions of reality, so the natural world is also thoroughly physical.

The problems facing Russellian monism are analogous to those facing neutral monism. Neutral monism claims that the fundamental entities are in themselves neither mental nor physical, but neutral: they can be accurately described as mental or physical if they satisfy certain extrinsic conditions. Some of these are conditions that qualify neutral events as mental and others that qualify them as physical, and it is possible for the same neutral event to satisfy conditions of both sorts. But events needn't be described in either mentalistic or physicalistic terms at all since they can all be exhaustively described and explained in terms of a conceptual framework that is neither mentalistic nor physicalistic.

One of the biggest challenges that has faced neutral monists is providing a clear account of what neutral entities are (Jaworski 2011: 260–2). Neutral entities are supposed to be entities that are in themselves neither mental nor physical, but that qualify as mental or physical if they satisfy certain conditions. Some neutral monists have been tempted to say little more than this. They say, for instance, that neutral entities are the ones that can factor into both physical and psychological laws, or that neutral entities are the entities that are capable of satisfying the conditions associated both with being mental and with being physical. The problem with these definitions of neutral entities is that they fail to provide an adequate means of identifying which entities in the world the neutral entities are, or of determining whether neutral entities even exist. By analogy, imagine that we arrive together at a meeting of an organization, and you want to know from me which person in the room is the organization's president. I respond: "Obviously the president is the person in the room who sets the meeting's agenda, who organizes the annual fundraiser, who was elected by the membership last January, who . . . ," and I continue to describe the relations that define the role the president occupies within the organization. You are doubtless disappointed by my response. The reason is that for all the information it

supplies, this response does not supply information that enables you to pick out which person in the room actually occupies the presidential role. Yet, that is precisely what you want: information that tells you which person in the room is the president. Attempts to define neutral entities simply in terms of their relations to mental and physical entities are analogous to my response, and seem to have an analogous shortcoming: they describe the role neutral entities are supposed to play within neutral monistic theory, relations to mental and physical entities that would be had by anything that qualified as a neutral entity, but they do not supply any information that would enable us to pick out which entities in the world, if any, actually play that role. If neutral monists cannot supply a definition that enables us to do this, then neutral monism remains merely an abstract possibility in logical space, not a genuine theory whose merits and demerits can be evaluated alongside other mind-body theories.

Not all neutral monists have rested content with purely relational definitions of the foregoing sort, but attempts to give more informative definitions have faced other problems. For instance, informative definitions of neutral entities have often had a mentalistic cast. James (1984a, b), for instance, characterized them as *pure experiences*. This generated the suspicion that for these neutral monists, so-called neutral entities were not really neutral but mental, and that they were not really neutral monists but closet idealists. Other attempts to provide informative definitions of neutral entities have had a physicalistic cast. Russell (1921), for instance, at one time characterized fundamental reality as an underlying *material* or *stuff* that was in itself neither mental nor physical. This characterization avoided the worry that neutral monism was a form of idealism, but it gave rise to the equal and opposite worry that neutral monism was really just a form of physicalism. Russell (1956) later abandoned this vocabulary, and in the end, seems to have favored describing the neutral entities simply as events. The category of events seems neutral enough. There is nothing built into the concept of an event that implies that events be mental or that they be physical. But this characterization of neutral entities borders on being uninformative in the way the relational characterizations of neutral entities discussed a moment ago are uninformative. What, after all, qualifies an event as neutral on this account? Granted, there is nothing built into the concept of an event that implies that events must be mental or physical, but by the same token, there is nothing built into the concept of an event that implies that events must be neutral; nothing implies, in other words, that they must be in themselves neither mental nor physical, but something else. Characterizing the basic entities simply as events falls short of specifying what qualifies any events that exist as neutral entities.

The demand for an informative characterization of neutral phenomena poses a serious challenge to neutral monism, and something analogous is true of Russellian monism. It is unclear what fundamental entities and protomental properties are supposed to be. Russellian monists could claim that the fundamental entities and protomental properties are precisely those entities and properties that in the right

combinations produce the full-blown mental properties that beings like you and I possess (Strawson 2006). In that case, though, Russellian monists face a problem analogous to the original problem facing neutral monists: their definition of the fundamental entities and protomental properties is insufficiently informative. In addition, Russellian monists face a related problem, the so-called 'combination problem,' the problem of explaining how protomental properties could combine to produce more sophisticated mental properties.

Hylomorphism sidesteps the foregoing issues entirely. Hylomorphists applaud the desire that motivates Russellian monism, the desire to situate mental phenomena more firmly in the natural world than brute laws of emergence by themselves allow. As we've seen, however, they propose to do this by taking structure to be a basic ontological and explanatory principle, and by taking thought, feeling, and perception to be structured activities. As a result, they needn't posit protomental properties, and they needn't explain how those properties (whatever they turn out to be) combine to form the thoughts, feelings, and perceptions beings like us possess.

Hylomorphism is no worse off, therefore, than its nonreductive naturalistic rivals, and in some respects, it is much better off than they are. The advantages that it enjoys over these theories provide yet another reason why hylomorphism deserves serious consideration.

14.4 Which Hylomorphism?

The kind of hylomorphic theory I've defended is not the only hylomorphic theory that has been advanced recently in the literature. Kit Fine (1999; 2008), Mark Johnston (2006), David Oderberg (2007), Kathryn Koslicki (2008), Michael Rea (2011), Anna Marmodoro (2013), and Robert Koons (2014) have also defended contemporary versions of hylomorphism. It will thus be helpful to say what reasons there are for preferring the kind of hylomorphism I've defended to theirs.

One reason to prefer the kind of hylomorphic theory I've defended is that the advantages of hylomorphism I described in Sections 14.1–14.3 pertain by and large only to it. If competing hylomorphic theories are able to provide elegant solutions to mind-body problems, their defenders have yet to demonstrate that.

In addition, it is possible to build a case in favor of the hylomorphic theory I've defended piecemeal, by tallying the various advantages it has over competing hylomorphic theories. Consider, for instance, the advantages it has over hylomorphic theories that claim that structures are parts, such as Fine's and Koslicki's. Johnston states the principal difficulty facing these structure-part views:

> A theorist might … [treat] a principle of unity as another sort of part, a property-part … On this conception of relations as property-parts, each material object will share parts with some state … So, if what it is for a certain cross to exist is for one plank to be attached to another at right angles then on the conception in question the cross consists of the planks and the relation

of being attached at right angles. But those same constituents will be parts of a certain state of the planks, namely, the state of the planks being attached together at right angles. This illustrates the central difficulty with the conception. The principle of unity for the state is distinct from the constituent relation of the state; the principle of unity in question requires that the constituent relation be instantiated by the constituent planks. But on the conception of objects as having their own principles of unity as parts, it will often be true—and this is fatal to the conception—that the principle of unity will not hold of itself and the other parts of the object. So it is with being attached together at right angles. This is a relation that could not possibly hold of three things, one of which is a relation . . . The relational properties that are the plausible "uniters" of ordinary parts into complex objects will not in general hold of the ordinary parts and the relational properties themselves. (2006: 672–3)

Structures are supposed to explain the unity of composite individuals and activities: they are what unite a complex thing's parts into a whole. Suppose, then, that *a* is a complex individual or activity with parts $b_1, b_2, \ldots, b_n$ that have structure *s*, and suppose that *s* is itself a part of *a* as structure-part theorists claim. Johnston's objection is that *s* must unite not only the *b*s to each other, but also *s* to the *b*s, yet in many cases *s* will not be a relation that can have itself as one of its own terms. In that case, however, the structure-part view will not have the kind of generality that is needed for a workable hylomorphic theory.

Structure-part theorists might respond by positing a further structure, s^*, which unites *s* to the *b*s. This would avoid Johnston's worry, but it would generate a different one, for the same problem that arose for *s* and the *b*s will now arise for s^*. In order for *s* and the *b*s to compose *a*, they have to be united to s^*. If s^* is what unites *s* and the *b*s, it remains unclear what unites s^* to *s* and the *b*s. Either it will be s^* itself, in which case structure-part theorists face Johnston's objection once again, or else it will be a further structure s^{**}, in which case structure-part theorists face the possibility of an infinite regress of structure-parts. The kind of hylomorphic view I've defended avoids these worries altogether since it denies that structures are parts, and this, I think, gives it an advantage over structure-part versions of hylomorphism.[2]

I think the hylomorphic view I've defended also has advantages over hylomorphic views that proliferate material objects, such as Fine's and Johnston's views.[3] Fine tells us,

[T]here [are] many more material objects than is commonly supposed. We are familiar with the prodigious ontology of the mereologists, according to which any occupied region of space-time, no matter how scattered or gerrymandered, will determine a material object. But this is nothing compared to the ontology of the present view. For to each such object of the mereologist, there will correspond a multitude of rigid embodiments . . . and a multitude of variable embodiments . . . [T]he flat unstructured objects of the mereologist represent a mere fraction of what there is. (1999: 73)

[2] Aristotle advances a similar argument at Metaphysics 1041b11–31.

[3] Oderberg (2014: 173–4) argues that Koslicki's view faces an analogous problem.

Similarly, according to Johnston,

> The material wholes that strike us as paradigmatically genuine are compact, well-articulated items that have self-maintaining principles of unity, items like Aristotle's favorite examples of substances, "the individual man and horse." But the least arbitrary form of the general theory of unity that is Hylomorphism tells us that besides these "genuine" wholes there is a vast plurality of wholes that ordinarily escape our attention. (2006: 698)

It is often taken to be a strike against a principle of unrestricted composition that it leads to a counterintuitive proliferation of the material objects that exist, not just things like tables, chairs, and people, but things like the object composed of my right foot and the Eiffel tower. If intuition provides any basis for rejecting unrestricted composition, it also provides a basis for rejecting theories like Fine's and Johnston's since they lead to a similar proliferation. Of course, we've seen that the hylomorphic theory I've defended faces the converse worry. Like van Inwagen's (1990) view, it leads to what many would consider the counterintuitive extermination of material objects such as tables, chairs, and planets (Section 6.2). We have seen, however, that the implications of the hylomorphic theory I've defended are not as extreme on this point as the implications of van Inwagen's view (Section 6.3), and that some version of the Denial is well motivated in any event (Section 7.4).

There is another, closely related reason to favor the hylomorphic view I've defended over the views of Fine and Johnston; namely, their views are committed to coincident objects. According to them, multiple objects can share the same constituent materials at the same time (Fine 1999: 73–4; 2008; Johnston 2006: 664–7). The alloy, for instance, is coincident with the piece of alloy, which is coincident with the statue. One worry about coincident objects is suggested by Fine himself:

> A scientist might explain the malleability of some alloy in terms of its chemical composition. But it is clear that we cannot expect to find a similar explanation of the malleability of the *piece* of the alloy . . . It is not as if further scientific investigation of the piece of alloy—as opposed to the alloy itself—might reveal it to be malleable. (2008: 103)

Recall that the kind of hylomorphism I've defended is committed to ontological naturalism, the claim that when it comes to determining what exists, empirical investigation is our best guide (Section 1.2). Many philosophers of a naturalistic stripe are likely to find a principle like ontological naturalism plausible in its own right quite independent of hylomorphism. For these philosophers, Fine's observation is likely to generate skepticism about the necessity and desirability of positing *piece* as a kind. If our best scientific descriptions, explanations, and methods reveal no difference between the alleged piece of alloy and the alloy itself, then the piece of alloy seems to be an unnecessary theoretical posit.

Naturally, hylomorphists and others who deny coincident objects face challenges of their own. Fine mentions two (2008: 106–7). First, they must provide descriptions of putative cases of coincidence that deny that there really are coinciding objects.

They must deny, for instance, either that the alloy exists, or that the piece of alloy exists, or that the statue exists. Second, if they countenance commonsense objects like the piece of alloy and the statue, those theories must account for the modal differences between the objects. For instance, they must account for why the alloy but not the statue can survive being squashed. We saw earlier that hylomorphists meet both of these challenges (Sections 6.2–6.3, 7.4). In some cases, they do this by denying that artifacts and natural bodies exist. If there are no statues and there are no pieces, then there is no statue or piece of alloy that shares all its parts with the alloy, and there are no coinciding objects whose differing modal properties need to be given an account. Hylomorphists of course countenance the existence of living things, so the same eliminativist strategy is not available to them when it comes to these things. We saw that in these cases they follow a different strategy (Sections 6.2–6.3). According to them, I am the only individual that is composed of the materials that are located exactly where I am. Those materials do not compose in addition a piece (or lump, or hunk, or chunk) of matter. The reason is that according to hylomorphists, composition happens only when a structured individual configures the materials that compose it, and I am the only individual located exactly where I am that is engaged in that configuring activity. It is true that located exactly where I am there are also numerous biofunctional parts such as my heart, brain, and kidneys, but the existence of these parts depends on my own existence. They enjoy a subordinate kind of unity which does not compromise my unity as a structured whole because it depends on the configuring activity of the whole (Section 7.1). The hylomorphic view thus provides a way of describing putative cases of coincidence that does not needlessly proliferate material objects.

There are also reasons to favor the hylomorphic theory I've defended over David Oderberg's (2007) and Anna Marmodoro's (2013). Their theories owe more to Thomistic philosophy than the approaches of other philosophers who have recently defended hylomorphic views. I have already had occasion to criticize the Thomistic view of parts that they endorse (Section 6.5). This by itself is enough, I think, to give the hylomorphic view I defend an edge. In addition, Oderberg endorses a variety of other traditional Thomistic doctrines, including the immateriality of the intellect. I criticized that doctrine in Section 8.3, along with the argument for it. E. J. Lowe (2010b) has argued, moreover, that Oderberg's view is problematic on account of its commitment to giving definitions *per genus et differentiam*, its commitment to the existence of prime matter, and its commitment to matter individuating substances. By contrast with Oderberg's view, the hylomorphic view I've defended is not committed to giving definitions *per genus et differentiam*, to the existence of prime matter, or to matter individuating substances.

Another liability of Oderberg's view is rhetorical: the traditional Thomistic-Aristotelian vocabulary he uses does more to obscure than to clarify the nature of hylomorphic metaphysics, as well as its similarities to and differences from other metaphysical frameworks, and its particular theoretical assets. Lowe's putative

criticism of hylomorphism illustrates the problems would-be hylomorphists are likely to encounter if they stick to a traditional vocabulary like Oderberg's:

> [Hylemorphism's] core difficulty lies in its central doctrine—that every concrete . . . individual substance, is a 'combination' of matter and form. For what . . . are we to *understand* by 'combination' in this sense? . . . [T]he supposed *need* to say something like this . . . supposedly arises in order to . . . [explain] how *a new substance* is brought into existence . . . rather than some previously existing things merely being rearranged . . . [S]ome types of 'rearrangement' are ontologically more weighty than others. When a free proton and a free electron are 'rearranged' by increasing the distance between them from one mile to two miles, there is no reason at all to suppose that a new concrete object is brought into existence. But when they are 'rearranged' so that the electron is captured by the proton and occupies an orbital around it, then indeed we have a new concrete object of a very different kind: a hydrogen atom. The object has certain features, notably certain *powers*, which are quite different from those of the protons and electrons . . . A new *form* is instantiated . . . namely, the form of *a hydrogen atom*. This form is the form of the newly created object, *the atom*, not that of the proton or the electron, nor even of the *pair* of them . . . The form does not, in any sense that I can understand, 'combine' with the proton and the electron so as to constitute, together with them, the atom. The only things that do any 'combining' are *the proton and the electron* . . . And the only things that *constitute* the atom are, again, the proton and the electron, which are its *parts* . . . I am perfectly happy to describe the case . . . in terms of 'combination' . . . 'constitution', and . . . 'form'. It's just that I don't need . . . the 'logical grammar' of the hylemorphist who uses these terms in his own distinctively technical fashion. (2011: 235–7)

Lowe does not advance a substantive objection to hylomorphism. In fact, we've seen that the hylomorphic theory I've defended would describe the case of the hydrogen atom in more or less the same way Lowe himself does (provided that hydrogen atoms really are structured wholes (Section 6.3)). Lowe's objection is instead to the 'logical grammar' of hylomorphic theories that adhere to a traditional Thomistic-Aristotelian vocabulary. Letting go of that vocabulary (in addition to any traditional doctrines that are no longer defensible) would enable would-be hylomorphists to achieve more effectively one of Oderberg's own objectives:

> If traditional metaphysics is to have a future . . . it must show itself to be a living system and method for doing philosophy. Its concepts must be deployed to tackle fundamental problems, including those that occupy contemporary thought. And it must shake itself free of the time-worn rhetoric that has, for several centuries, been used to vilify it . . . by virtue of a highly defective understanding of just what the concepts and theses of neo-Aristotelianism actually mean. (2007: x)

Another competing hylomorphic theory is defended by Michael Rea (2011). His hylomorphic theory is closer to mine than the others I've described so far. He endorses a substance-attribute ontology that takes structures to be properties, and that takes properties to be powers. There are nevertheless important differences between his theory and the one I defend which I think favor the latter.

First, Rea says two things which suggest that he favors a pure dispositionalist account of powers like Shoemaker's (1980). When describing the challenges that dispositional monists face, the two he mentions are specific to pure dispositionalism (2011: 346): the always packing never traveling objection, and the worries about grounding (Section 4.6). In addition, he rejects the act-potency distinction on the grounds that a dispositionalist view posits dispositions "all the way down" (348). Yet, the identity theory of powers is capable of positing dispositions "all the way down" without rejecting the act-potency distinction (Section 4.1). He thus appears to reject the identity theory. Since the only dispositional monist alternative is pure dispositionalism, he appears tacitly committed to the latter. I think that this commitment is a liability for reasons discussed in Section 4.6: there are well-rehearsed objections to pure dispositionalism which the identity theory of powers avoids. I think this is one factor that gives the hylomorphic theory I defend an edge.

Second, Rea seems to be committed to properties being abundant. He says, for instance, that the natures of substances are perfectly natural properties in Lewis' (1983b) sense (347). The idea that there are properties that are less than perfectly natural, however, suggests that properties are abundant, and we saw in Chapter 2 that there are good reasons to reject abundant properties, especially when it comes to discussions in the philosophy of mind.

Third, Rea denies that properties are either universals or particulars (343–4). He instead takes them to be primitive powers with some characteristics that are typically attributed to particulars and others that are typically attributed to universals. Like a particular, for instance, he says that a power can have a location and enter into causal relations, but like a universal, the same power can have multiple locations (346). While Rea sees the postulation of primitive powers as an asset, others might see it as a liability, especially those who are convinced that there are good reasons to posit universals or tropes. Moreover, Rea is motivated to reject a formulation in terms of either universals or particulars by arguments that purport to show that the universal-particular distinction cannot be successfully drawn (Ramsey 1925; MacBride 2005). But it is possible to formulate the trope theory defended in Chapter 3 without having to invoke a universal-particular distinction. That formulation would provide examples of properties (this apple's redness, that person's bravery) and would simply deny that these properties were entities of the sort universals are supposed to be: it would deny that they were repeatable, that numerically one property was had in common by numerically distinct individuals, and so on. Trope theorists are not obliged to offer a positive characterization of particulars that goes beyond this. At least, they are not obliged to say more about particulars than Rea says about primitive powers, and in that sense, a theory committed to tropes, like the one I've defended, is no worse off than a theory committed to primitive powers like Rea's.[4] In fact, it might

[4] Marmodoro (2013) claims that Rea's view cannot explain similarity—how, for instance, apple *a*'s redness manages to be similar to apple *b*'s redness. It is difficult to see how this conclusion follows for Rea

be better off since tropes are more familiar to metaphysicians than Rea's primitive powers, and some philosophers might favor an ontology with familiar entities over an ontology with newfangled ones.

Fourth, Rea posits a primitive notion of numerical sameness that is not identity (345–6). Like the appeal to the primitive powers, some philosophers might consider this departure from the familiar notion of identity a cost.

Fifth, Rea suggests that substances can be reduced to or identified with powers (344, 355 n. 9). The result appears to be a bundle theory of substance similar to the bundle theories endorsed by some trope theorists (Stout 1921, 1923; Williams 1953, 1986; Campbell 1981, 1983, 1990; Simons 1994; Mumford 2012). The worry here is that there are well-rehearsed objections to trope bundles (Martin 1980; Armstrong 1989a: 114–16; Molnar 2003: 47–51). Granted, Rea's substances are bundles of primitive powers instead of bundles of tropes, but his powers are similar enough to tropes that it's not implausible to suppose that objections could be advanced against his view that are analogous to those advanced against trope bundles.

Sixth, Rea's hylomorphism faces a difficulty diagnosed by Koons (2014), which centers on Rea's account of what it means for a structured whole to unite the powers of its diverse parts. Rea states his account as follows:

A power p_0 of an object x *unites* distinct powers $p_1, \ldots, p_n =_{\text{df}}$

(i) p_0 is intrinsic to x,

(ii) each of $p_1, \ldots, p_n$ is a nature of at least one of x's parts,

(iii) p_0 is grounded in or identical to a certain sort of cooperative manifestation CM of $p_1, \ldots, p_n$,

(iv) every power intrinsic to x that is at least partly grounded in CM is identical with, reducible to, or at least partly grounded in p_0, and

(v) there is no power intrinsic to x that is distinct from both p_0 and CM and that grounds p_0. (2011: 349)

According to Koons, there is a tension among Rea's claims here. On the one hand, composite wholes have natures. This implies that their powers are fundamental, for on Rea's view, natures are fundamental powers. On the other hand, composite wholes are also supposed to unite the powers of their parts. Given (i)–(v), this seems to imply that the powers of a whole must all be grounded in the powers of its parts, but this implies that those powers are not fundamental. As a result, says Koons, "There is an important gap in Rea's account... he hasn't shown that substances have unique natures" (2014: 159). The gap can only be filled, Koons suggests, by allowing that at least some of the powers of the parts of a composite whole are grounded in the nature of the whole itself. The hylomorphic theory I've defended makes precisely that allowance. In that sense, it can also be said to have an advantage over Rea's.

for the same reason it is difficult to see how it follows for a trope theory of the sort I've defended (Section 3.3).

Finally, it is worth mentioning that the hylomorphic theory I defend avoids what Rea takes to be a danger for hylomorphic theories: that of "disconnecting our metaphysics of material objects from empirical reality" (2011: 342). Because the hylomorphic theory I defend rounds out the picture of structure by appeal to empirical sources, it avoids this worry.

What of Koons' (2014) own hylomorphic theory? Among the theories I've mentioned here, his is probably closest to my own. According to Koons, a hylomorphic theory worthy of the name has to meet two conditions. First, it has to ensure that the persistence of a composite whole is grounded in the coordinated operation of its parts. Second, it has to ensure that the whole cannot act or be acted upon except through (at least in part) the powers of its parts. Koons argues that his hylomorphic theory satisfies these conditions. That theory centers on a thesis he calls *parts as sustaining instruments* (PASI), which is based on the following theses:

> *Upward Sustenance*: The persistence and cooperation of the substance's parts *cause* the persistence of the whole.
>
> *Upward Power Migration*: Some (or all) causal powers migrate from parts to the whole.
>
> *Teleological Subordination*: The powers and activities of the parts are teleologically ordered to some end pertaining to the whole.
>
> *Parts are sustaining instruments*: The constraints of Upward Sustenance, Upward Power Migration, and Teleological Subordination are all jointly realized.

According to Koons, Upward Sustenance involves viewing substantial forms (what I have been calling 'individual-making structures') as processes. Upward Power Migration implies that elementary particles have powers that they manifest in different ways under different conditions. In particular, as Koons understands the thesis, they manifest powers in one way when they are parts of composite wholes, and they manifest powers in other ways when they are 'in the wild', apart from composite wholes. Finally, according to Teleological Subordination, the powers that elementary particles manifest when they are parts of a composite whole are ordered to the natural ends of the organism.

According to Koons, PASI satisfies his two conditions for an adequate hylomorphic theory. The first condition is satisfied because PASI implies the following thesis:

> *Sustenance*: For any composite substance x with proper parts the yy's and any moment t at which the substance exists, the existence of x at t is wholly grounded in the actual persistence of some process P in some interval of time beginning at some instant t_0 and ending at t, which process P is such that its participants from t_0 until and including t are exactly the yy's (or exactly x itself and the yy's).

The second condition is satisfied because PASI implies the following thesis:

> *Instrumentation*: For any composite substance x, any causal power P of x at any moment t, there is a proper part y of x at t, a power P^* of y at t, such that P^* is at least partly grounded in P, and the exercise of P^* at t would contribute to the natural end of x.

I've said that among the hylomorphic theories we've considered, Koons' PASI account is most similar to mine. This is easy to appreciate. Upward Sustenance corresponds to the claim that individual-making structures are activities in which a structured whole engages. We saw in Section 6.3 that hylomorphists of my stripe needn't insist on the term 'activities'; Koons' 'processes' is fine as well. Likewise, Upward Power Migration corresponds to the claim that things like the electron gain new properties when they are incorporated into structured wholes (Section 6.3). Physical materials do not lose the powers they possess when they are 'in the wild' (Section 8.2); their powers are instead limited or directed by the wholes in which they are incorporated since those wholes have additional parts whose powers inhibit, enhance, or otherwise influence the powers of other parts (Section 6.5). Koons appears to agree with something like this picture (167–70). Something's nature cannot be altered, he says. Consequently, if an electron's nature comprises an array of powers, it cannot lose any of those powers without ceasing to exist. Rather, the electron's nature comprises an array of powers, some of which it manifests only when incorporated into a structured whole, others of which it manifests only when it is 'in the wild,' and yet others of which it manifests in both cases. Because of this, a composite whole is able to harness the powers of its parts in ways that contribute to its own biological ends, a claim that corresponds to Koons' Teleological Subordination.

The hylomorphic view I've defended is thus similar in its implications to Koons' PASI account, and like that account, it implies theses that are similar to Sustenance and Instrumentation. A structured whole cannot exist apart from the materials that it configures, and it exists exactly as long as its configuring activity (or process) continues (Section 6.1). Moreover, all the powers of a structured whole are essentially embodied in the powers of its parts (Sections 8.2–8.3), a claim which corresponds to Koons' Instrumentation.

What, then, should we say about Koons' version of hylomorphism? Insofar as Koons' claims are compatible with the hylomorphic view I've defended, I do not take his version of hylomorphism to be a competitor. I take it instead to be a more abstract statement of the kind of hylomorphic theory I've looked to defend.

The considerations I've advanced in this section do not by any means provide knockdown arguments against competing hylomorphic theories, but they do provide some reason for preferring the hylomorphic theory I've defended to its competitors.

14.5 Conclusion

The goal of this chapter was to argue that hylomorphism should be taken seriously—at least as seriously as more familiar theories in the philosophy of mind such as nonreductive physicalism, emergentism, and Russellian monism. To that end, I have argued that hylomorphism elegantly solves a number of philosophical problems in metaphysics and the philosophy of mind. It also has strong intuitive credentials. It meshes with commonsense intuitions about structure, and also with the intuitions a number of philosophers have about the notion of organization. If there is a difference between the way hylomorphists approach talk of organization or structure and the way other authors do, it is that hylomorphists take seriously the ontological demands that talk of organization or structure seems to imply. Hylomorphism is committed to ontological naturalism, the claim that when it comes to determining what exists, empirical investigation is our best guide. Since many of our best empirical descriptions, explanations, and methods appear to posit various kinds of organization or structure, those descriptions, explanations, and methods give us prima facie reason to think that organization or structure exists.

Structure realism takes empirical appeals to structure at face value. Structure, it says, is a basic ontological and explanatory principle: descriptions, explanations, and methods that posit structure cannot in general be reduced to, or paraphrased, or eliminated in favor of nonstructural descriptions and explanations. This straightforward approach to structure is the one favored by hylomorphists. There are alternatives which claim that it is possible in principle to describe and explain everything without appeal to structure. These include structure eliminativism, structure reductivism, and nonreductive structure antirealism. There are nevertheless reasons for preferring structure realism.

First, structure realism is the simplest, most direct explanation for why appeals to structure in biology, neuroscience, and other scientific disciplines are successful. It should thus operate as our default position in the absence of compelling reasons to do otherwise—reasons it is the burden of structure realism's opponents to supply. Second, structure realism is preferable to structure eliminativism because unlike the latter, it is able to explain the effectiveness of scientific appeals to structure. In addition, it is evidentially easier to be a structure realist than a structure reductivist or a nonreductive structure antirealist. Structure reductivists have to show that the structures we posit in our empirical descriptions and explanations can be identified with things that can be exhaustively described and explained without appeal to structure, and nonreductive structure antirealists must show that all references to structure can be paraphrased in ways that have no implicit or explicit commitment to it. Both of these are very ambitious projects. Hylomorphists do not have an a priori argument that these projects cannot be carried out, but what gives their theory an edge, they insist, is that they needn't carry out any analogous project. In addition,

nonreductive structure antirealism has trouble explaining how scientific appeals to structure manage to be successful if fundamentally everything is unstructured. Nonreductive structure antirealists claim that we use structural predicates and terms because they enable us to satisfy peculiar descriptive and explanatory interests, but just because appeals to structure satisfy some of our interests does not guarantee that those appeals will manage to be successful at describing or explaining anything. Moreover, it is unclear what explains the fact that we have these interests. Surely, say hylomorphists, the best explanation is that we are interested in describing and explaining the real behavior of things, and that behavior involves various kinds of structure.

A general argument against structure realism appeals to Ockham's razor. But ontological parsimony only becomes a decisive factor in theory choice when competing theories are all coherent, all consistent with the empirical data, and equal in explanatory power, and according to hylomorphists, opponents of structure have yet to establish that their theories are consistent with the empirical data, and have explanatory power to rival a theory that endorses structure. In addition, rival theories have complexities that offset their greater ontological parsimony. There are thus reasons to find structure realism preferable to the alternatives. This provides a further reason to take hylomorphism seriously.

Hylomorphism also enjoys advantages over competing mind-body theories. For instance, it solves the major problems emergentists face: the problem of emergence and the problem of downward causation. And there are similar reasons to prefer it to nonreductive physicalism. There are, moreover, reasons to prefer it to physicalist theories in general. For one thing, it is evidentially easier to be a hylomorphist than a physicalist since hylomorphism, unlike physicalism, does not make a totalizing claim about the descriptive and explanatory power of a single branch of science.

Hylomorphism also enjoys advantages over Russellian monism, which claims that fundamental physical entities have categorical mental properties that combine to produce mental properties like the ones you and I have. Russellian monists have difficulty accounting for how protomental properties could combine to produce higher-level mental properties, and they also face a problem giving an informative characterization of what protomental properties are supposed to be. Hylomorphism sidesteps these issues entirely.

Finally, among the various hylomorphic theories that have been advanced in recent years, there are reasons to prefer the hylomorphic theory I've defended to its competitors. First, it avoids the problems confronting hylomorphic theories like Kit Fine's and Kathryn Koslicki's, which take structures to be parts. In addition, it avoids problems with coincident objects and does not lead to an embarrassing proliferation of material objects like the hylomorphic theories endorsed by Fine and Mark Johnston. Nor does it carry the heavy Thomistic commitments of David Oderberg's and Anna Marmodoro's theories—commitments which prove to be

liabilities. Finally, there are reasons to prefer it to Michael Rea's hylomorphic view, which, among other things, posits a number of unfamiliar primitives.

None of these are knockdown arguments in favor of the kind of hylomorphism I've articulated, but I think they are sufficient to show that it deserves serious consideration.

References

Ackrill, J. L. 1972–3. Aristotle's Definitions of *Psuchê*, *Proceedings of the Aristotelian Society* 73: 1991–3. Reprinted in *Articles on Aristotle, Volume 4: Psychology and Aesthetics*, eds. Jonathan Barnes, Malcolm Schofield, and Richard Sorabji. Bristol: Classics Press, 1979, 65–75.

Adler, Mortimer. 1993. *The Difference of Man and the Difference It Makes*. New York: Fordham University Press.

Alexander, Samuel. 1920. *Space, Time, and Deity*. New York: Macmillan.

Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.

Armstrong, D. M. 1978a. *Realism and Nominalism: Universals and Scientific Realism*, Vol. 1. Cambridge: Cambridge University Press.

Armstrong, D. M. 1978b. *A Theory of Universals: Universals and Scientific Realism*, Vol. 2. Cambridge: Cambridge University Press.

Armstrong, D. M. 1983. *What Is a Law of Nature?* Cambridge: Cambridge University Press.

Armstrong, D. M. 1989a. *Universals: An Opinionate Introduction*. Boulder, CO: Westview Press.

Armstrong, D. M. 1989b. *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press.

Armstrong, D. M. 1996a. Dispositions as Categorical States. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 15–18.

Armstrong, D. M. 1996b. Place's and Armstrong's Views Compared and Contrasted. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 33–48.

Armstrong, D. M. 1996c. Reply to Martin. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 88–104.

Armstrong, D. M. 1997. *A World of States of Affairs*. Cambridge: Cambridge University Press.

Armstrong, D. M. 1999. The Causal Theory of Properties: Properties according to Shoemaker, Ellis, and Others. *Philosophical Topics* 26: 25–37.

Armstrong, D. M. 2002. Two Problems for Essentialism. In Ellis 2002, 167–71.

Armstrong, D. M. 2004a. *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Armstrong, D. M. 2004b. How Do Particulars Stand to Universals? *Oxford Studies in Metaphysics* 1: 139–54.

Armstrong, D. M. 2005. Four Disputes about Properties. *Synthese* 144: 309–20.

Armstrong, D. M. 2010. *Sketch for a Systematic Metaphysics*. Oxford: Oxford University Press.

Bacon, John. 1995. *Universals and Property Instances: The Alphabet of Being*. Oxford: Blackwell Publishers.

Baker, Lynne Rudder. 2000. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.

Bauer, William A. 2012. Four Theories of Pure Dispositions. In *Properties, Powers and Structures: Issues in the Metaphysics of Realism*, eds. Alexander Bird, Brian Ellis, and Howard Sankey. New York: Routledge, 139–62.

Bechtel, William. 2007. Reducing Psychology while Maintaining Its Autonomy via Mechanistic Explanations. In *The Matter of the Mind*, eds. Maurice Schouten and Huib Looren de Jong. Chichester: Blackwell Publishing, 172–98.

Bechtel, William. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge.

Bechtel, William, and Robert C. Richardson. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press.

Belnap, Nuel, and Thomas Steel. 1976. *The Logic of Questions*. New Haven, CT: Yale University Press.

Bennett, Jonathan. 1988. *Events and Their Names*. Indianapolis, IN: Hackett Publishing.

Bickle, John. 2003. *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer Academic Publishers.

Bird, Alexander. 1998. Dispositions and Antidotes. *Philosophical Quarterly* 48: 227–34.

Bird, Alexander. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.

Black, Robert. 2000. Against Quidditism. *Australasian Journal of Philosophy* 78: 87–104.

Blackburn, Simon. 1990. Filling in Space. *Analysis* 50: 62–5.

Block, Ned, ed. 1980. *Readings in Philosophy of Psychology*, 2 vols. Cambridge, MA: Harvard University Press.

Block, Ned, and Jerry Fodor. 1972. What Psychological States Are Not. *Philosophical Review* 80: 159–81. Reprinted with revisions by the authors in Block 1980, 237–50.

Bradley, F. H. 1899. *Appearance and Reality: A Metaphysical Essay*. New York: Macmillan.

Bradley, M. C. 1979. Two Logical Connection Arguments and Some Principles about Causal Connection. *Erkenntnis* 14: 1–23.

Broad, C. D. 1925. *The Mind and Its Place in Nature*. London: Routledge and Kegan Paul.

Burge, Tyler. 1993. Mind-Body Causation and Explanatory Practice. In Burge 2007, 344–62.

Burge, Tyler. 2006. Postscript to "Mind-Body Causation and Explanatory Practice". In Burge 2007, 363–82.

Burge, Tyler. 2007. *Foundations of Mind*. Oxford: Oxford University Press.

Burke, Michael B. 1994. Dion and Theon: An Essential Solution to an Ancient Puzzle. *Journal of Philosophy* 91: 129–39.

Burke, Michael B. 1996. Coinciding Objects: Reply to Lowe and Denkel. *Analysis* 57: 11–18.

Callus, Daniel A. 1961. The Origins of the Problem of the Unity of Form. In *The Dignity of Science: Studies in the Philosophy of Science Presented to William Humbert Kane, O.P.* Washington, DC: Thomist Press.

Camazine, Scott, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz, and Eric Bonabeau. 2001. *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Cameron, J. R. 1999. Plural Reference. *Ratio* 12: 128–47.

Campbell, Keith. 1976. *Metaphysics: An Introduction*. Encino, CA: Dickenson Publishing.

Campbell, Keith. 1981. The Metaphysic of Abstract Particulars. *Midwest Studies in Philosophy* 6: 477–86. Reprinted in Laurence and Macdonald, eds. 1998, 277–94.

Campbell, Keith. 1983. Abstract Particulars and the Philosophy of Mind. *Australasian Journal of Philosophy* 61: 129–41.

Campbell, Keith. 1990. *Abstract Particulars*. Oxford: Basil Blackwell.

Campbell, Neil A. 1996. *Biology*, 4th Edition. San Francisco, CA: Benjamin Cummings Publishing.

Campbell, Neil A., Jane B. Reece, and Lawrence G. Mitchell. 1999. *Biology*, 5th Edition. San Francisco, CA: Benjamin Cummings.

Carruthers, Peter. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.

Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Causey, Robert L. 1977. *Unity of Science*. Boston, MA: D. Reidel Publishing.

Chalmers, David J. 1996. *The Conscious Mind*. Oxford: Oxford University Press.

Chalmers, David J. 2002a. Consciousness and Its Place in Nature. In *Philosophy of Mind*, ed. David J. Chalmers. Oxford: Oxford University Press, 247–72.

Chalmers, David J. 2002b. Does Conceivability Entail Possibility? In *Conceivability and Possibility*, eds. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press, 145–200.

Chalmers, David J. 2004. Epistemic Two-Dimensional Semantics. *Philosophical Studies* 118: 153–226.

Chalmers, David J. 2009. The Two-Dimensional Argument against Materialism. In *Oxford Handbook of Philosophy of Mind*, ed. Brian McLaughlin. Oxford: Oxford University Press, 313–35.

Chalmers, David J. 2010. *The Character of Consciousness*. Oxford: Oxford University Press.

Chisholm, Roderick M. 1957. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press.

Chisholm, Roderick M. 1970. Events and Propositions. *Nous* 4: 15–24.

Churchland, Patricia S. 1986. *Neurophilosophy*. Cambridge, MA: MIT Press.

Churchland, Paul M. 1981. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press/Bradford Books.

Cottingham, Robert Stoothoff, and Dugald Murdoch. 1984. *The Philosophical Writings of Descartes*, Vol. II. Cambridge: Cambridge University Press.

Crane, Tim. 1994. Physicalism (2): Against Physicalism. In *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan. Oxford: Blackwell Publishers, 479–84.

Crane, Tim and D. H. Mellor. 1990. There Is No Question of Physicalism. *Mind* 99: 185–206.

Craver, Carl F. 2002. Structures of Scientific Theories. In *The Blackwell Guide to the Philosophy of Science*, eds. Peter Machamer and Michael Silberstein. Oxford: Blackwell Publishers, 55–79.

Craver, Carl F. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. New York: Oxford University Press.

Cummins, Robert. 1975. Functional Analysis. *Journal of Philosophy* 72: 741–64.

Daly, Chris. 1994. Tropes. *Proceedings of the Aristotelian Society*, New Series 94: 253–61.

Davidson, Donald. 1963. Actions, Reasons, and Causes. *Journal of Philosophy* 60: 685–700. Reprinted in Davidson 1980, 3–19.

Davidson, Donald. 1969. The Individuation of Events. In *Essays in Honor of Carl G. Hempel*, ed. N. Rescher. Dordrecht: D. Reidel, 216–34. Reprinted in Davidson 1980, 163–80.

Davidson, Donald. 1970. Mental Events. In *Experience and Theory*, eds. L. Foster and J. W. Swanson. Amherst, MA: University of Massachusetts Press. Reprinted in Davidson 1980, 207–25.

Davidson, Donald. 1974. Psychology as Philosophy. In *Philosophy of Psychology*, ed. S. C. Brown. London: Macmillan. Reprinted in Davidson 1980, 229–39.

Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.

Dennett, Daniel C. 1978. *Brainstorms*. Cambridge, MA: Bradford Books.

Dennett, Daniel C. 1991. Real Patterns. *Journal of Philosophy* 88: 27–51.

Dennett, Daniel C. 1995. The Unimagined Preposterousness of Zombies. *Journal of Consciousness Studies* 2: 322–6.

Dewey, John. 1958. *Experience and Nature*. New York: Dover Publications.

Donagan, Alan. 1963. Universals and Metaphysical Realism. *Monist* 47: 211–46.

Dretske, Fred. 1977. Laws of Nature. *Philosophy of Science* 44: 248–68.

Dretske, Fred. 1988. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press/Bradford Books.

Dretske, Fred. 1991. Dretske's Replies. In *Dretske and His Critics*, ed. Brian McLaughlin. Oxford: Basil Blackwell, 180–221.

Dupré, John. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Egan, Andy. 2004. Second-Order Predication and the Metaphysics of Properties. *Australasian Journal of Philosophy* 82: 48–66.

Ekman, Paul. 2007. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 2nd Edition. New York: Owl Books.

Ellis, Brian. 1999. Causal Powers and Laws of Nature. In *Causation and Laws of Nature*, ed. H. Sankey. Dordrecht: Kluwer Academic Publishers, 77–97.

Ellis, Brian. 2001. *Scientific Essentialism*. Cambridge: Cambridge University Press.

Ellis, Brian. 2002. *The Philosophy of Nature: A Guide to the New Essentialism*. Chesham: Acumen.

Ellis, Brian, and Caroline Lierse. 1994. The Fundamental Importance of Natural Kinds. La Trobe University: Victorian Centre for the History and Philosophy of Science Preprint Series 3/94.

Engelhard, Kristina. 2010. Categories and the Ontology of Powers: A Vindication of the Identity Theory of Powers. In Marmodoro 2010, 41–57.

Evans, Gareth. 1978. Can There Be Vague Objects? *Analysis* 38: 208.

Evans, Gareth. 1982. *The Varieties of Reference*, ed. John McDowell. Oxford: Clarendon Press.

Feigl, Herbert. 1958. The Mental and the Physical. In *Minnesota Studies in the Philosophy of Science*, Vol. 2. Minneapolis, MN: University of Minnesota Press, 370–497.

Field, Hartry. 1972. Tarski's Theory of Truth. *Journal of Philosophy* 69: 347–75.

Fine, Kit. 1999. Things and Their Parts. *Midwest Studies in Philosophy* 23: 61–74.

Fine, Kit. 2008. Form and Coincidence. *Proceedings of the Aristotelian Society, Supplementary Volume* 82: 101–18.

Fodor, Jerry. 1968a. *Psychological Explanation: An Introduction to the Philosophy of Psychology*. New York: Random House.

Fodor, Jerry. 1968b. The Appeal to Tacit Knowledge in Psychological Explanation. *Journal of Philosophy* 65: 627–40.

Fodor, Jerry. 1974. Special Sciences, or The Disunity of Science as a Working Hypothesis. *Synthese* 28: 97–115.

Foster, John. 1982. *The Case for Idealism*. London: Routledge and Kegan Paul.

Furth, Montgomery. 1978. Transtemporal Stability in Aristotelian Substances. *Journal of Philosophy* 75: 624–46.

Garber, Daniel. 1985. Leibniz and the Foundations of Physics: The Middle Years. In *The Natural Philosophy of Leibniz*, eds. Kathleen Okruhlik and James Robert Brown. Dordrecht: D. Reidel, 27–130.

Garcia, Robert K. 2015. Is Trope Theory a Divided House? In *The Problem of Universals in Contemporary Philosophy*, eds. Gabrielle Galluzzo and Michael Loux. Cambridge: Cambridge University Press, 133–55.

Garcia, Robert K. Forthcoming. Two Ways to Particularize a Property. *Journal of the American Philosophical Association*.

Geach, Peter. 1967. *Mental Acts*. New York: Humanities Press.

Geach, Peter. 1969. *God and the Soul*. London: Routledge and Kegan Paul.

Gendler, Tamar Szabó, and John Hawthorne, eds. 2002. *Conceivability and Possibility*. Oxford: Oxford University Press.

Glennan, Stuart. 1996. Mechanisms and the Nature of Causation. *Erkenntnis* 44: 50–71.

Glennan, Stuart. 2002. Rethinking Mechanistic Explanation. *Philosophy of Science* 69: S342–53.

Goldman, Alvin. 1970. *A Theory of Human Action*. Princeton, NJ: Princeton University Press.

Goodman, Nelson. 1956. A World of Individuals. In *The Problem of Universals: A Symposium*, eds. I. M. Bochenski, Alonzo Church, and Nelson Goodman. Notre Dame, IN: University of Notre Dame Press, 15–31.

Goodman, Nelson. 1966. *The Structure of Appearance*, 2nd Edition. Indianapolis, IN: Bobbs–Merrill Co., Inc.

Goodman, Nelson, and W. V. Quine. 1947. Steps toward a Constructive Nominalism. *Journal of Symbolic Logic* 12: 105–22.

Grene, Marjorie. 1963. *A Portrait of Aristotle*. Chicago, IL: University of Chicago Press.

Grice, H. P. 1989. *Studies in the Ways of Words*. Cambridge: Harvard University Press.

Grimes, Thomas R. 1988. The Myth of Supervenience. *Pacific Philosophical Quarterly* 69: 152–60.

Groff, Ruth, and John Greco, eds. 2012. *Powers and Capacities in Philosophy: The New Aristotelianism*. London: Routledge.

Hacking, Ian. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.

Haldane, J. B. S. 1947. *What Is Life?* New York: Boni and Gaer.

Hamlyn, D. W. 1953. Behaviour. *Philosophy* 28: 132–45.

Hamlyn, D. W. 1968. *Aristotle's* De Anima, Books II, III (with Certain Passages from Book I). Oxford: Oxford University Press.

Harré, R., and E. H. Madden. 1975. *Causal Powers: A Theory of Natural Necessity*. Oxford: Basil Blackwell.

Hartman, Edwin. 1977. *Substance, Body, and Soul: Aristotelian Investigations*. Princeton, NJ: Princeton University Press.

Haugeland, John. 1998. *Having Thought*. Cambridge, MA: Harvard University Press.

Hawthorne, John. 2001. Causal Structuralism. *Philosophical Perspectives* 15: 361–78.

Heil, John. 2003. *From an Ontological Point of View*. Oxford: Clarendon Press.

Heil, John. 2005. Dispositions. *Synthese* 144: 343–56.

Heil, John, and Alfred Mele, eds. 1993. *Mental Causation*. Oxford: Clarendon Press.

Hempel, Carl. 1969. Reduction: Ontological and Linguistic Facets. In *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, eds. Sydney Morgenbesser, Patrick Suppes, and Morton White. New York: St Martin's Press, 179–99.

Hochberg, Herbert. 2001a. *The Positivist and the Ontologist: Bergmann, Carnap, and Logical Realism*. Amsterdam: Editions Rodopi B.V.

Hochberg, Herbert. 2001b. *Russell, Moore and Wittgenstein: The Revival of Realism*. Berlin: Ontos-Verlag.

Honderich, Ted. 1982. The Argument for Anomalous Monism. *Analysis* 42: 59–64.

Honderich, Ted. 1988. *A Theory of Determinism*. Oxford: Oxford University Press.

Hooker, Clifford A. 1981. Towards a General Theory of Reduction. *Dialogue* 20: 38–60, 201–35, 496–529.

Hopson, James A. 1975. The Evolution of Cranial Display Structures in Hadrosaurian Dinosaurs. *Paleobiology* 1: 21–43.

Irwin, T. H. 1988. *Aristotle's First Principles*. Oxford: Clarendon Press.

Jacobs, Jonathan. 2011. Power Qualities, Not Pure Powers. *Monist* 94: 81–102.

Jaeger, Werner. 1934. *Aristotle: Fundamentals of the History of His Development*, trans. Richard Robinson. Oxford: Oxford University Press.

James, William. 1984a. A World of Pure Experience. In *William James: The Essential Writings*, ed. Bruce W. Wilshire. Albany, NY: State University of New York Press, 178–97.

James, William. 1984b. Does 'Consciousness' Exist? In *William James: The Essential Writings*, ed. Bruce W. Wilshire. Albany, NY: State University of New York Press, 162–77.

Jaworski, William. 2009. The Logic of How-Questions. *Synthese* 166: 133–55.

Jaworski, William. 2011. *Philosophy of Mind: A Comprehensive Introduction*. Chichester: Wiley-Blackwell.

Jaworski, William. 2012. Powers, Structures, and Minds. In *Powers and Capacities in Philosophy: The New Aristotelianism*, eds. Ruth Groff and John Greco. New York: Routledge, 145–71.

Johnson, W. E. 1964. *Logic*. New York: Dover Publications.

Johnston, Mark. 1992. How to Speak of the Colors. *Philosophical Studies* 68: 221–63.

Johnston, Mark. 2006. Hylomorphism. *Journal of Philosophy* 103: 652–98.

Kahn, Charles H. 1992. Aristotle on Thinking. In *Essays on Aristotle's* De Anima, eds. M. Nussbaum and A. O. Rorty. Oxford: Clarendon Press, 359–80.

Kauffman, Stuart A. 1971. Articulation of Parts Explanation in Biology and the Rational Search for Them. In *PSA 1970*, eds. R. C. Buck and R. S. Cohen. Dordrecht: D. Reidel.

Kim, Jaegwon. 1973. Causation, Nomic Subsumption and the Concept of Event. *Journal of Philosophy* 70: 217–36. Reprinted in Kim 1993a, 3–21.

Kim, Jaegwon. 1976. Events as Property Exemplifications. In *Action Theory*, eds. Myles Brand and Douglas Walton. Dordrecht: D. Reidel, 159–77. Reprinted in Kim 1993a, 33–52.

Kim, Jaegwon. 1984. Concepts of Supervenience. *Philosophy and Phenomenological Research* 45: 153–76. Reprinted in Kim 1993a, 53–78.

Kim, Jaegwon. 1987. 'Strong' and 'Global' Supervenience Revisted. *Philosophy and Phenomenological Research* 48: 315–26. Reprinted in Kim 1993a, 79–91.

Kim, Jaegwon. 1988. Supervenience for Multiple Domains. *Philosophical Topics* 16: 129–50. Reprinted in Kim 1993a, 109–30.

Kim, Jaegwon. 1989. The Myth of Nonreductive Physicalism. *Proceedings and Addresses of the American Philosophical Association* 63: 31–47. Reprinted in Kim 1993a, 265–84.

Kim, Jaegwon. 1990. Supervenience as a Philosophical Concept. *Metaphilosophy* 21: 1–27. Reprinted in Kim 1993a, 131–60.

Kim, Jaegwon. 1991. Dretske on How Reasons Explain Behavior. In *Dretske and His Critics*, ed. Brian McLaughlin. Oxford: Basil Blackwell, 52–72. Reprinted in Kim 1993a, 285–308.

Kim, Jaegwon. 1992. Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research* 52: 1–26. Reprinted in Kim 1993a, 309–35.

Kim, Jaegwon. 1993a. *Supervenience and Mind*. Cambridge: Cambridge University Press.

Kim, Jaegwon. 1993b. The Non-Reductivist's Troubles with Mental Causation. In Heil and Mele, eds., 189–210. Reprinted in Kim 1993a, 336–57.

Kim, Jaegwon. 1998. *Mind in a Physical World*. Cambridge, MA: MIT Press/Bradford Books.

Kim, Jaegwon. 2003. Supervenience, Emergence, Realization, Reduction. In *The Oxford Handbook of Metaphysics*, eds. Michael J. Loux and Dean W. Zimmerman. Oxford: Oxford University Press, 556–84.

Kim, Jaegwon. 2005. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.

Kim, Jaegwon. 2006a. *Philosophy of Mind*, 2nd Edition. Boulder, CO: Westview Press.

Kim, Jaegwon. 2006b. Emergence: Core Ideas and Issues. *Synthese* 151: 547–59.

Kingsley, Robert E. 2000. *Concise Text of Neuroscience*, 2nd Edition. Philadelphia, PA: Lippincott Williams and Wilkins.

Kitcher, Philip. 1984. 1953 and All That: A Tale of Two Sciences. *Philosophical Review* 93: 335–73.

Koons, Robert. 2014. Staunch vs. Faint-Hearted Hylomorphism: Toward an Aristotelian Account of Composition. *Res Philosophica* 91: 151–78.

Koslicki, Kathrin. 2008. *The Structure of Objects*. Oxford: Oxford University Press.

Kripke, Saul A. 1972. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Ladyman, James, and Don Ross. 2007. *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Laurence, Stephen, and Cynthia Macdonald, eds. 1998. *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Blackwell Publishers.

Leftow, Brian. 2001. Souls Dipped in Dust. In *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, ed. Kevin Corcoran. Ithaca, NY: Cornell University Press, 120–38.

Levinson, Jerrold. 1978. Properties and Related Entities. *Philosophy and Phenomenological Research* 39: 102–15.

Lewis, David K. 1966. An Argument for the Identity Theory. *Journal of Philosophy* 63: 17–25. Reprinted in Lewis 1983a, 99–107.

Lewis, David K. 1968. Counterpart Theory and Quantified Modal Logic. *Journal of Philosophy* 65: 113–26. Reprinted in Lewis 1983a, 26–39.

Lewis, David K. 1970. How to Define Theoretical Terms. *Journal of Philosophy* 67: 427–46. Reprinted in Lewis 1983a, 78–95.

Lewis, David K. 1971. Counterparts of Persons and Their Bodies. *Journal of Philosophy* 68: 203–11. Reprinted in Lewis 1983a, 47–54.

Lewis, David K. 1972. Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy* 50: 249–58. Reprinted in Lewis 1999, 248–61.

Lewis, David K. 1973a. *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, David K. 1973b. Causation. *Journal of Philosophy* 70: 556–67. Reprinted in Lewis 1986b, 159–72.

Lewis, David K. 1983a. *Philosophical Papers*, Vol. 1. Oxford: Oxford University Press.

Lewis, David K. 1983b. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61: 343–77. Reprinted in Lewis 1999: 8–55.

Lewis, David K. 1986a. *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Lewis, David K. 1986b. *Philosophical Papers*, Vol. 2. Oxford: Oxford University Press.

Lewis, David K. 1986c. Causal Explanation. In Lewis 1986b, 214–40.

Lewis, David K. 1986d. Against Structural Universals. *Australasian Journal of Philosophy* 64: 25–46. Reprinted in Lewis 1999: 78–110.

Lewis, David K. 1991. *Parts of Classes*. Oxford: Basil Blackwell.

Lewis, David K. 1994. Lewis, David: Reduction of Mind. In *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan. Oxford: Blackwell Publishers, 412–31.

Lewis, David K. 1997. Finkish Dispositions. *Philosophical Quarterly* 47: 143–58.

Lewis, David K. 1999. *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.

Lewis, David K. 2008. Ramseyan Humility. In *Conceptual Analysis and Philosophical Naturalism*, eds. David Braddon-Mitchell and Robert Nola. Oxford: Oxford University Press, 203–22.

Loar, Brian. 1990. Phenomenal States. *Philosophical Perspectives* 4, ed. James E. Tomberlin. Atascadero, CA: Ridgeview Publishing, 81–108.

Locke, John. 1959. *An Essay Concerning Human Understanding*, 2 vols. New York: Dover Publications.

Lombard, Lawrence Brian. 1986. *Events: A Metaphysical Study*. London: Routledge and Kegan Paul.

Lombard, Lawrence Brian. 1998. Ontologies of Events. In Laurence and Macdonald, eds. 1998, 277–94.

Loux, Michael J. 2006. *Metaphysics: A Contemporary Introduction*, 3rd Edition. New York: Routledge.

Loux, Michael J. 2008. *Primary Ousia: An Essay on Aristotle's* Metaphysics *Z and H*. Ithaca, NY: Cornell University Press.

Loux, Michael J. 2015. An Exercise in Constituent Ontology. In *The Problem of Universals in Contemporary Philosophy*, eds. Gabrielle Galluzzo and Michael Loux. Cambridge: Cambridge University Press, 9–45.

Lowe, E. J. 1996. *Subjects of Experience*. Cambridge: Cambridge University Press.

Lowe, E. J. 2006. *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Clarendon Press.

Lowe, E. J. 2010a. On the Individuation of Powers. In Marmodoro 2010, 1–26.

Lowe, E. J. 2010b. David S. Oderberg's *Real Essentialism*. *Philosophical Quarterly* 60: 648–52.

Lowe, E. J. 2011. A Neo-Aristotelian Substance Ontology. In *Contemporary Aristotelian Hylomorphism*, ed. Tuomas E. Tahko. Cambridge: Cambridge University Press.

Lycan, William G. 1987. *Consciousness*. Cambridge, MA: MIT Press.

Lycan, William G. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.

MacBride, Fraser. 2005. The Universal-Particular Distinction: A Dogma of Metaphysics? *Mind* 114: 565–614.

Macdonald, Cynthia. 1998. Tropes and Other Things. In Laurence and Macdonald, eds. 1998, 229–350.

Machamer, Peter, Lidley Darden, and Carl F. Craver. 2000. Thinking about Mechanisms. *Philosophy of Science* 67: 1–25.

Malcolm, Norman. 1968. The Conceivability of Mechanism. *Philosophical Review* 77: 45–72.

Manley, David. 2002. Properties and Resemblance Classes. *Nous* 36: 75–96.

Marcus, Eric. 2004. Why Zombies Are Inconceivable. *Australasian Journal of Philosophy* 82: 477–90.

Marmodoro, Anna. 2010. *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge.

Marmodoro, Anna. 2013. Aristotelian Hylomorphism without Reconditioning. In *Philosophical Inquiry* 36: 5–22.

Marmodoro, Anna. Forthcoming. *Aristotelian Powers at Work: Reciprocity without Symmetry in Causation*. In *Causal Powers*, ed. Jonathan Jacobs. Oxford: Oxford University Press.

Martin, C. B. 1980. Substance Substantiated. *Australasian Journal of Philosophy* 58: 3–10.

Martin, C. B. 1993. Power for Realists. In *Ontology, Causality and Mind: Essays in Honor of D. M. Armstrong*, eds. John Bacon, Keith Campbell, and Lloyd Reinhart. Cambridge: Cambridge University Press, 175–86.

Martin, C. B. 1994. Dispositions and Conditionals. *Philosophical Quarterly* 44: 1–8.

Martin, C. B. 1996a. Properties and Dispositions. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 71–87.

Martin, C. B. 1996b. Replies to Armstrong and Place. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 126–46.

Martin, C. B. 1997. On the Need for Properties: The Road to Pythagoreanism and Back. *Synthese* 112: 193–231.

Martin, C. B. 2007. *The Mind in Nature*. Oxford: Oxford University Press.

Martin, C. B., and John Heil. 1998. Rules and Powers. *Philosophical Perspectives* 12: 238–312.

Martin, C. B., and John Heil. 1999. The Ontological Turn. *Midwest Studies in Philosophy* 23: 34–60.

Martin, C. B., and Karl Pfeifer. 1986. Intentionality and the Non-Psychological. *Philosophy and Phenomenological Research* 46: 531–54.

Mayr, Ernst. 1982. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Belknap Press of Harvard University.

Mayr, Ernst. 1997. *This Is Biology: The Science of the Living World*. Cambridge, MA: Belknap Press of Harvard University.

McKitrick, Jennifer. 2003. The Bare Metaphysical Possibility of Bare Dispositions. *Philosophy and Phenomenological Research* 66: 349–69.

McKitrick, Jennifer. 2010. Manifestations as Effects. In Marmodoro 2010, 73–83.

McLaughlin, Brian P. 1992. The Rise and Fall of British Emergentism. In *Emergence or Reduction?: Essays in the Prospects of Nonreductive Physicalism*, eds. Ansgar Beckerman, Hans Flohr, and Jaegwon Kim. Berlin: Walter de Gruyter, 49–93.

McLaughlin, Brian P. 1995. Varieties of Supervenience. In *Supervenience: New Essays*, eds. Elias Savellos and Umit D. Yalcin. Cambridge: Cambridge University Press, 16–59.

Melden, A. I. 1961. *Free Action*. London: Routledge and Kegan Paul.

Melnyk, Andrew. 1997. How to Keep the 'Physical' in Physicalism. *Journal of Philosophy* 94: 622–37.

Melnyk, Andrew. 2003. *A Physicalist Manifesto*. New York: Cambridge University Press.

Meltzoff, A. N. 1995. Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18–Month-Old Children. *Developmental Psychology* 31: 838–50.

Merleau-Ponty, Maurice. 2002. *Phenomenology of Perception*, 2nd Edition. New York: Routledge.

Merricks, Trenton. 2001a. *Objects and Persons*. Oxford: Clarendon Press.

Merricks, Trenton. 2001b. How to Live Forever without Saving Your Soul. In *Soul, Body, and Survival*, ed. Kevin Corcoran. Ithaca, NY: Cornell University Press, 183–200.

Merricks, Trenton. 2005. Composition and Vagueness. *Mind* 114: 615–37.

Merricks, Trenton. 2007. Remarks on Vagueness and Arbitrariness. *Mind* 116: 115–19.

Mertz, D. W. 1996. *Moderate Realism and Its Logic*. New Haven, CT: Yale University Press.

Mill, J. S. 1843. *A System of Logic*. London: John W. Parker and Son, West Strand.

Miller, Fred D. 2012. Aristotle on the Separability of Mind. In *The Oxford Handbook of Aristotle*, ed. Christopher Shields. Oxford: Oxford University Press.

Miller, Jonathan. 1978. *The Body in Question*. New York: Random House.

Molnar, George. 2003. *Powers: A Study in Metaphysics*, ed. Stephen Mumford. Oxford: Oxford University Press.

Montague, Richard. 1960. On the Nature of Certain Philosophical Entities. *Monist* 53: 159–94. Reprinted in *Formal Philosophy: Selected Papers of Richard Montague*, ed. Richmond H. Thomason. New Haven, CT: Yale University Press, 1974, 148–87.

Montero, Barbara. 1999. The Body Problem. *Nous* 33: 183–200.

Mumford, Stephen. 1998. *Dispositions*. Oxford: Oxford University Press.

Mumford, Stephen. 1999. Intentionality and the Physical: A New Theory of Disposition Ascription. *Philosophical Quarterly* 49: 215–25.

Mumford, Stephen. 2004. *Laws in Nature*. London: Routledge.

Mumford, Stephen. 2006. The Ungrounded Argument. *Synthese* 149: 471–89.

Mumford, Stephen. 2012. The Power of Power. In Groff and Greco, eds., 9–24.

Mumford, Stephen, and Rani Lill Anjum. 2011. *Getting Causes from Powers*. Oxford: Oxford University Press.

Nagel, Ernest. 1961. *The Structure of Science*. Indianapolis, IN: Hackett Publishing.

Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.

Nagel, Thomas. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. Oxford: Oxford University Press.

Nersessian, Nancy J. 1992. How Do Scientists Think? Capturing the Dynamics of Conceptual Change. In *Cognitive Models of Science: Minnesota Studies in the Philosophy of Science*, Vol. 15, ed. Ronald N. Giere. Minneapolis, MN: University of Minnesota Press, 3–44.

Nickles, Thomas. 1975. Two Concepts of Intertheoretic Reduction. *Journal of Philosophy* 70: 181–201.

Noë, Alva. 2004. *Action in Perception*. Cambridge, MA: MIT Press.

Nussbaum, Martha. 1984. Aristotelian Dualism: Reply to Howard Robinson. *Oxford Studies in Ancient Philosophy* 2: 197–207.

Nussbaum, Martha C., and Hilary Putnam. 1992. Changing Aristotle's Mind. In *Essays on Aristotle's* De Anima, eds. M. Nussbaum and A. O. Rorty. Oxford: Clarendon Press, 27–56.

Nussbaum, Martha C., and Amélie Oksenberg Rorty, eds. 1992. *Essays on Aristotle's* De Anima. Oxford: Clarendon Press.

O'Connor, Timothy, and Jonathan D. Jacobs. 2003. Emergent Individuals. *Philosophical Quarterly* 53: 540–55.

Oddie, Graham. 1982. Armstrong on the Eleatic Principle and Abstract Entities. *Philosophical Studies* 41: 285–95.

Oderberg, David S. 2007. *Real Essentialism*. New York: Routledge.

Oderberg, David S. 2014. Is Form Structure? In *Neo-Aristotelian Perspectives in Metaphysics*, eds. Daniel Novotný and Lukáš Novak. New York: Routledge.

Oliver, Alex. 1994. Are Subclasses Parts of Classes? *Analysis* 54: 215–23.

Olson, Eric. T. 1995. Why I Have No Hands. *Theoria* 61: 182–97.

Olson, Eric T. 1996. *The Human Animal: Personal Identity without Psychology*. Oxford: Oxford University Press.

Oppenheim, Paul, and Hilary Putnam. 1958. Unity of Science as a Working Hypothesis. In *Minnesota Studies in the Philosophy of Science* 2: 3–36.

Papineau, David. 1993. *Philosophical Naturalism*. Oxford: Blackwell Publishers.

Paull, R. Cranston, and Theodore Sider. 1992. In Defense of Global Supervenience. *Philosophy and Phenomenological Research* 52: 833–54.

Peacocke, Christopher. 1989. No Resting Place: A Critical Notice of the View from Nowhere. *Philosophical Review* 98: 65–83.

Penfield, Wilder. 1975. *The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain*. Princeton, NJ: Princeton University Press.

Pickel, Bryan, and Nicholas Mantegani. 2012. A Quinean Critique of Ostrich Nominalism. *Philosophers' Imprint* 12: 1–21.

Place, U. T. 1996a. Intentionality as the Mark of the Dispositional. *Dialectica* 50: 91–120.

Place, U. T. 1996b. Dispositions as Intentional States. In *Dispositions: A Debate*, ed. Tim Crane. London: Routledge, 19–32.

Poland, Jeffrey. 1994. *Physicalism: The Philosophical Foundations*. New York: Oxford University Press.

Prasada, Sandeep, Krag Ferenz, and Todd Haskell. 2002. Conceiving of Entities as Objects and Stuff. *Cognition* 83: 141–65.

Price, H. H. 1953a. *Thinking and Experience*. London: Hutchinson and Co.

Price, H. H. 1953b. Universals and Resemblance. From chapter 1 of *Thinking and Experience* (London: Hutchinson University Library). Reprinted in *Metaphysics: Contemporary Readings*, ed. Michael J. Loux. London: Routledge, 2001, 20–41.

Prior, A. N. 1949. Determinables, Determinates and Determinants. *Mind* 58: 178–94.

Prior, Elizabeth W., Robert Pargetter, and Frank Jackson. 1982. Three Theses about Dispositions. *American Philosophical Quarterly* 19: 251–7.

Psillos, Stathis. 2006. What Do Powers Do When They Are Not Manifested? *Philosophy and Phenomenological Research* 72: 137–55.

Putnam, Hilary. 1963. Brains and Behavior. In *Analytical Philosophy Second Series*, ed. R. J. Butler. Oxford: Basil Blackwell and Mott, 1–19. Reprinted in Putnam 1975b, 325–41.

Putnam, Hilary. 1967. Psychological Predicates. In *Art, Mind, and Religion*, eds. W. H. Capitan and D. D. Merrill. Pittsburgh, PA: University of Pittsburgh Press, 37–48. Reprinted as 'The Nature of Mental States' in Putnam 1975b, 429–40.

Putnam, Hilary. 1970. On Properties. In *Essays in Honor of Carl G. Hempel*, ed. N. Rescher. Dordrecht: D. Reidel. Reprinted in Putnam 1975a, 305–22.

Putnam, Hilary. 1975a. *Mathematics, Matter, and Method: Philosophical Papers*, Vol. 1. New York: Cambridge University Press.

Putnam, Hilary. 1975b. *Mind, Language, and Reality: Philosophical Papers*, Vol. 2. Cambridge: Cambridge University Press.

Putnam, Hilary. 1975c. The Meaning of 'Meaning'. In *Language, Mind and Knowledge: Minnesota Studies in the Philosophy of Science*, Vol. 7, ed. Keith Gunderson. Minneapolis, MN: University of Minnesota Press. Reprinted in Putnam 1975b, 215–71.

Putnam, Hilary. 1975d. What Is Mathematical Truth? In Putnam 1975a, 60–78.

Quine, W. V. 1948. On What There Is. *Review of Metaphysics* 2: 21–38. Reprinted in *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953, 1–19.

Quine, W. V. 1953. *From a Logical Point of View*. Cambridge, MA: Harvard University Press.

Quine, W. V. 1960. *Word and Object*. Cambridge, MA: MIT Press.

Quinton, Anthony. 1957. Properties and Classes. *Proceedings of the Aristotelian Society* 48: 33–58.

Ramsey, F. P. 1925. Universals. *Mind* 34: 401–17. Reprinted in *F. P. Ramsey: Philosophical Papers*, ed. D. H. Mellor. Cambridge: Cambridge University Press, 1990, 8–30.

Rea, Michael C. 2011. Hylomorphism Reconditioned. *Philosophical Perspectives* 25: 341–58.

Robinson, Howard. 1982. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge: Cambridge University Press.

Robinson, Howard. 2014. Modern Hylomorphism and the Reality and Causal Power of Structure: A Skeptical Investigation. *Res Philosophica* 91: 203–14.

Rosenthal, David M. 2005. *Consciousness and Mind*. New York: Oxford University Press.

Ross, W. D. 1957. The Development of Aristotle's Thought. *Proceedings of the British Academy* 43: 63–78.

Ruse, Michael. 2001. *Can a Darwinian Be a Christian? The Relationship between Science and Religion*. Cambridge: Cambridge University Press.

Russell, Bertrand. 1912. *The Problems of Philosophy*. London: Home University Library.

Russell, Bertrand. 1921. *The Analysis of Mind*. London: Unwin Paperbacks.

Russell, Bertrand. 1956. Mind and Matter. In *Portraits from Memory and Other Essays*. New York: Simon and Schuster, 145–65.

Ryle, Gilbert. 1949. *The Concept of Mind*. Abingdon: Routledge.

Salmon, Nathan. 1981. *Reference and Essence*. Princeton, NJ: Princeton University Press.

Salmon, Wesley C. 1989. Four Decades of Scientific Explanation. In *Minnesota Studies in the Philosophy of Science*, Vol. 13, eds. Philip Kitcher and Wesley C. Salmon. Minneapolis, MN: University of Minnesota Press.

Schaffer, Jonathan. 2009. On What Grounds What. In *Metametaphysics: New Essays on the Foundations of Ontology*, eds. David Chalmers, David Manley, and Ryan Wasserman. Oxford: Oxford University Press, 347–83.

Schaffner, Kenneth. 1967. Approaches to Reduction. *Philosophy of Science* 34: 137–47.

Searle, John R. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

Searle, John R. 2004. *Mind: A Brief Introduction*. Oxford: Oxford University Press.

Sellars, Wilfrid. 1957. Substance and Form in Aristotle. *Journal of Philosophy* 54: 688–99.

Sellars, Wilfrid. 1963a. Empiricism and the Philosophy of Mind. In *Science, Perception and Reality*. London: Routledge and Kegan Paul, 127–96.

Sellars, Wilfrid. 1963b. Abstract Entities. *Review of Metaphysics* 16: 627–71.

Shields, Christopher. 2014. *Aristotle*, 2nd Edition. London: Routledge.

Shoemaker, Sydney. 1980. Causality and Properties. In *Time and Cause: Essays Presented to Richard Taylor*, ed. Peter van Inwagen. Dordrecht: D. Reidel, 109–35. Reprinted in *Identity, Cause and Mind*, expanded edition. Oxford: Oxford University Press, 2003, 206–33.

Sider, Theodore. 1993. Van Inwagen and the Possibility of Gunk. *Analysis* 53: 285–9.

Sider, Theodore. 2001. *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford: Clarendon Press.

Sider, Theodore. 2012. *Writing the Book of the World*. Oxford: Oxford University Press.

Simon, Herbert A. 1969. *Sciences of the Artificial*. Cambridge, MA: MIT Press.

Simons, Peter. 1994. Particulars in Particular Clothing: Three Trope Theories of Substance. *Philosophy and Phenomenological Research* 54: 553–75. Reprinted in Laurence and Macdonald, eds., 1998, 364–84.

Simpson, George Gaylord. 1964. *This View of Life: The World of an Evolutionist*. New York: Harcourt, Brace, and World.

Sisko, John E. 1999. On Separating the Intellect from the Body: Aristotle's *De Anima* III.4, 429a20–b5. *Archiv für Geschichte der Philosophie* 81: 249–67.

Sklar, Lawrence. 1967. Types of Inter-Theoretic Reduction. *British Journal for the Philosophy of Science* 18: 109–24.

Smith, Justin Erik Halldór. 2002. Leibniz's Hylomorphic Monad. *History of Philosophy Quarterly* 19: 21–42.

Sober, Elliott. 1985. Panglossian Functionalism and the Philosophy of Mind. *Synthese* 64: 165–93.

Sommerhoff, Gerd. 1969. The Abstract Characteristics of Living Systems. In *Systems Thinking: Selected Readings*, ed. F. E. Emery. Harmondsworth: Penguin, 147–202.

Sorabji, Richard. 1974. Body and Soul in Aristotle. *Philosophy* 49: 63–8.

Sorabji, Richard. 1992. Intentionality and Physiological Processes: Aristotle's Theory of Sense-Perception. In Nussbaum and Rorty, eds., 195–226.

Sorensen, Roy. 2002. The Art of the Impossible. In *Conceivability and Possibility*, eds. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press, 337–68.

Sperry, Roger. 1984. Emergence. In *The Omni Interviews*, ed. Pamela Weintraub. New York: Ticknor and Fields, 187–207.

Spirtes, Peter, Clark Glymour, and Richard Scheines. 1993. *Causation, Prediction and Search*. New York: Springer-Verlag.

Stoljar, Daniel. 2010. *Physicalism*. New York: Routledge.

Stout, G. F. 1921. *The Nature of Universals and Propositions*. Oxford: Oxford University Press.

Stout, G. F. 1923. Symposium: Are the Characteristics of Particular Things Universal or Particular? Part II. *Proceedings of the Aristotelian Society, Supplementary Volume* 3: 114–22.

Strawson, Galen. 2006. Realistic Monism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies* 13: 3–31.

Strawson, P. F. 1974. *Subject and Predicate in Logic and Grammar*. London: Methuen.

Suppe, Frederick. 1974. The Search for Philosophic Understanding of Scientific Theories. In *The Structure of Scientific Theories*, ed. Frederick Suppe. Chicago, IL: University of Illinois Press, 3–232.

Swinburne, Richard G. 1980. A Reply to Shoemaker. In *Applications of Inductive Logic*, eds. L. J. Cohen and M. Hesse. Oxford: Clarendon Press, 316–17.

Swoyer, Christopher. 1982. The Nature of Natural Laws. *Australasian Journal of Philosophy* 60: 203–23.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tooley, Michael. 1977. The Nature of Laws. *Canadian Journal of Philosophy* 7: 667–98.

Trumpler, Maria. 1997. Converging Images: Techniques of Intervention and Forms of Representation of Sodium-Channel Proteins in Nerve Cell Membranes. *Journal of the History of Biology* 30: 55–89.

Van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Van Fraassen, Bas. 1989. *Laws and Symmetry*. Oxford: Oxford University Press.

Van Inwagen, Peter. 1981. The Doctrine of Arbitrary Undetached Parts. *Pacific Philosophical Quarterly* 62: 123–7.

Van Inwagen, Peter. 1990. *Material Beings*. Ithaca, NY: Cornell University Press.

Van Inwagen, Peter. 1994. Composition as Identity. *Philosophical Perspectives* 8: 207–20.

Van Inwagen, Peter. 2002. *Metaphysics*, 2nd Edition. Boulder, CO: Westview Press.

Wiggins, David. 1968. On Being in the Same Place at the Same Time. *Philosophical Review* 77: 90–5.

Wilkes, Kathleen V. 1978. *Physicalism*. Atlantic Highlands, NJ: Humanities Press.

Williams, Bernard. 1986. Hylomorphism. *Oxford Studies in Philosophy* 4. Oxford: Oxford University Press, 189–99. Reprinted in *The Sense of the Past: Essays in the History of Philosophy* by Bernard Williams. Princeton, NJ: Princeton University Press, 2006.

Williams, Donald C. 1953. The Elements of Being. *Review of Metaphysics* 7: 3–18. Reprinted in *Metaphysics: Contemporary Readings*, ed. Michael J. Loux. London: Routledge, 2001, 57–64.

Williams, Donald C. 1986. Universals and Existents. *Australasian Journal of Philosophy* 64: 1–14.

Wilson, Jessica. 2005. Supervenience-Based Formulations of Physicalism. *Nous* 39: 426–59.

Wilson, Jessica. 2010. What Is Hume's Dictum, and Why Believe It? *Philosophy and Phenomenological Research* 80: 595–637.

Wimsatt, William C. 1974. Complexity and Organization. In *PSA 1972* (*Boston Studies in the Philosophy of Science*, Vol. 2), eds. K. F. Schaffner and R. S. Cohen. Dordrecht: D. Reidel, 67–86.

Wimsatt, William C. 1985. Forms of Aggregativity. In *Human Nature and Natural Knowledge*, eds. Alan Donagan, Anthony N. Perovich, Jr., and Michael F. Wedin. Dordrecht: D. Reidel, 259–93.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, trans. G. E. M. Anscombe. New York: Macmillan.

Wolterstorff, Nicholas. 1960. Qualities. *Philosophical Review* 69: 183–200.

Wolterstorff, Nicholas. 1970. *On Universals: An Essay in Ontology*. Chicago, IL: University of Chicago Press.

Worrall, John. 1989. Structural Realism: The Best of Both Worlds? *Dialectica* 43: 99–124.

Yablo, Stephen. 1992. Mental Causation. *Philosophical Review* 101: 245–80.

Young, J. Z. 1971. *An Introduction to the Study of Man*. Oxford: Clarendon Press.

Ziff, Paul. 1958. About Behaviourism. *Analysis* 18: 132–6.

Zimmerman, Dean. 2003. Material People. In *The Oxford Handbook of Metaphysics*, eds. Michael J. Loux and Dean W. Zimmerman. Oxford: Oxford University Press, 491–526.

Index